DEPARTMENT OF LEARNING, INFORMATICS, MANAGEMENT AND ETHICS

Karolinska Institutet, Stockholm, Sweden

NURSE STAFFING LEVELS, CARE LEFT UNDONE, & PATIENT MORTALITY IN ACUTE HOSPITALS

Jane E Ball

Karolinska Institutet

Stockholm 2017

NURSE STAFFING LEVELS, CARE LEFT UNDONE, & PATIENT MORTALITY IN ACUTE HOSPITALS

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Jane E Ball

Principal Supervisor:

Professor Carol Tishelman
Karolinska Institutet
Department of LIME
Medical Management Centre

Co-supervisor(s):

Professor Peter Griffiths
University of Southampton
Faculty of Health Sciences
& Karolinska Institutet
Foreign Adjunct Professor
Department of LIME

Dr Rikard Lindqvist
Karolinska Institutet
Department of LIME
Medical Management Centre

Opponent:

Professor Jack Needleman
University of California, Los Angeles
Department of Health Policy & Management
Fielding School of Public Health

Examination Board:

Professor Kjell Asplund (Coordinator)
Umeå University
Department of Public Health & Clinical Medicine

Professor Ulrica von Thiele Schwarz
Karolinska Institutet
Department of LIME
Medical Management Centre

Professor Christine Brulin
Umeå University
Department of Nursing

"A person may cause evil to others not only by his actions but by his inaction, and in either case he is justly accountable to them for the injury"

John Stuart Mill, 1859

ABSTRACT

Failures in the care provided in hospitals may threaten patient safety. Research since the 1980s has established an association between lower registered nurse (RN) staffing levels and higher risk of death in acute hospitals. When RNs have more patients to care for, there is also a greater chance that some nursing care is missed. Yet despite the volume of research undertaken we have little empirical evidence about how or why RN staffing impacts on mortality and studies have rarely directly considered the contribution of medical and support staffing alongside RNs.

The research in this thesis has investigated the relationships between RN staffing levels, nursing care that is left undone (also termed 'missed care'), and patient risk of death in general acute hospitals. Using an observational design with cross-sectional data, the research examined RN staffing and mortality, concurrently with medical and support worker staffing, before going on to study the part played by care left undone as a potential consequence of nurse staffing, other factors associated with care left undone, and care left undone as a predictor of patient mortality.

The EU funded study 'RN4Cast' was the primary data source used with additional analyses based on routinely reported national data. Nurse staffing, nurse-rated quality and safety, and care left undone were assessed through a survey of RNs on medical and surgical wards. Mortality and patient risk factors were identified from administrative data sources.

In Study I better ward-based RN staffing levels were associated with a lower risk of mortality for medical patients, in regression models controlling for both medical and support worker staffing for the 31 RN4Cast NHS hospital trusts in England (RR 0.89, p=0.001). There was no difference in support worker staffing and risk of medical patient mortality. Higher support worker levels were associated with poorer outcomes for surgical patients (RR 1.01, p = 0.053).

Necessary nursing care was reported as being missed due to lack of time by 86% of RNs surveyed in England and 74% in Sweden. (Study II, Study III). Lower levels of missed care were associated with more positively rated quality of care and patient safety environment. (Study II). When RN staffing levels were lower, the risk of care being missed was increased. (Study II, Study III). However, higher levels of support workers were not associated with less care being left undone by RNs in England. (Study II). Despite some differences in the care context between England and Sweden, the same fundamental relationship between care left undone and RN staffing was found in both countries, and in an analysis of data from nine EU countries. (Study II, III, & IV).

Care left undone was significantly associated with mortality following common surgical procedures (OR 1.16, CI 1.04-1.29) in Study IV, which examined case-mix adjusted 30-day mortality in 300 hospitals in nine countries. Care left undone mediated the relationship between RN staffing and patient mortality following surgery. Care left undone offers a plausible causal pathway between RN staffing levels and patient mortality. (Study IV).

This is the first study to demonstrate that nursing left undone is on the causal path between low staffing levels and higher mortality. Taking this research together with the studies that have preceded it, this study provides evidence that higher RN staffing levels is causally linked to reduction in the risk of fatal harm to patients.

LIST OF SCIENTIFIC PAPERS

I.   Griffiths P, Ball J, Murrells T, Jones S, Rafferty AM. Registered nurse, healthcare support worker, medical staffing levels and mortality in English hospital trusts: a cross-sectional study. BMJ Open. 2016 Jan 1;6(2):e008751.

II. Ball J E, Murrells T, Rafferty AM, Morrow E, Griffiths P. 'Care left undone' during nursing shifts: associations with workload and perceived quality of care. BMJ Quality & Safety. 2014;23(2):116-25.

III. Ball J E, Griffiths P, Rafferty AM, Lindqvist R, Murrells T, Tishelman C. A Cross-sectional study of 'care left undone' on nursing shifts in hospitals. Journal of Advanced Nursing 72(9): 2086–2097 doi: 10.1111/jan.12976

IV. Ball J E, Bruyneel L, Aiken L, Sermeus W, Sloane D, Rafferty AM, Lindqvist R, Tishelman C, Griffiths P. Post-operative mortality, missed care and nurse staffing in nine countries. *Submitted for publication*.

CONTENTS

LIST OF ABBREVIATIONS & KEY TERMS

Acute hospital Trusts	NHS Trusts (see below) which include one or several hospital sites at which general acute care is provided.
'Care left undone'	Nursing care that has been missed or not completed fully. Also referred to as 'missed care'. As defined in this research: necessary nursing care activities that have been left undone due to lack of time.
HCSW	Health Care Support Worker – a generic term to refer to staff who have lower levels of qualification or health care preparation than Registered Nurses (RNs), who deliver some nursing care and are part of the nursing team. Also referred to as 'nursing aids', 'nursing assistants', 'nursing associates', 'nursing auxiliaries'.
ITU/ICU	Intensive Therapy Unit/Intensive Care Unit (synonymous).
NICE	National Institute for Health and Care Excellence (England).
NHPPD	Nursing hours per patient day.
NHS	National Health Service.
NHS Trust	The governance body for an NHS organisation providing health care services.
Nurse	Where the term 'nurse' is used without further explanation, it is as a shorthand for Registered Nurse (RN).
Nursing staff	Used to refer to the entire team of staff, including registered nurses (RNs), and nursing support staff/health care support workers (HCSWs). Physicians and others categories of staff are not included in the term.
Nursing support staff	Staff who have lower levels of qualification or health care preparation than RNs, who deliver some nursing care and are included in the ward nursing team. This staff group includes: 'health care support workers', 'health care assistants', 'nursing aids', 'nursing assistants', 'nursing associates', 'nursing auxiliaries'.
RN	Registered Nurse (with a minimum of 3 years nursing education).
RN4Cast	Acronym for the RN Forecasting Study (funded by EU 7th framework).

PROLOGUE

When I told my brother that my PhD was on whether nurse staffing levels affect quality of care in hospitals, he answered directly: 'Of course they do. Next question please'.

He isn't alone in finding my choice of research topic almost absurdly self-evident. It stands to reason that the number of nurses on duty will affect nurses' ability to deliver all the care needed by patients. And that this is likely to affect the quality of care patients receive, which in turn may impact on their eventual outcomes. Decades of research would also broadly support this view: hospitals with better registered nurse (RN) staffing levels have better patient outcomes. When I have discussed the research with nurses, or written about it in blogs, the reaction sometimes verges on outrage: 'We've known this for years - why waste more time and money researching something so obvious?'

Why indeed. Peppered throughout the paragraph above are: 'seems obvious', 'stands to reason', and 'likely'. But is 'it makes good sense' strong enough grounds for health care managers to base practice on, or ministers to base policies on? With finite resources, and increasing demands for health care, questions about how best to staff a hospital repeatedly arise. The decisions taken, locally or nationally, need to be informed by the best possible evidence. Whilst there has been much research in this field, it is an evolving landscape. The broad sweep of the terrain is clear, but there remain some uncertainties, some areas less well-mapped in the knowledge base.

It's been a long while since I was a clinical nurse juggling competing demands on a busy medical ward. Fighting down the sense of panic, with a growing realisation that what needs doing is more than I can get done. Hoping that somehow, by luck or judgment, the amount of harm caused by what I leave undone is minimal. Leaving a shift with a nagging doubt, and finding it increasingly difficult to imagine staying in nursing.

But I found my place in nursing. It's here. In this research. In this dissertation: dissecting and analysing the factors that shape workloads, that influence whether care is left undone, and the consequences for patient safety of missed nursing care. Failing to have the right number of nursing staff on duty with sufficient skill levels has consequences – for nurses, the care they provide and the patients who entrust themselves into the care of a hospital. Research has a part to play in unpicking these relationships, and by sharing the emerging findings we can hope to help decision-makers uphold Florence Nightingale's first principle and abide by the Hippocratic oath: to do the patient no harm.

1 INTRODUCTION

Health service providers around the world face similar tensions between managing pressures to contain costs, at the same time as meeting rising demands for health care services, whilst maintaining or ideally improving quality. Nursing is a key component of hospital care, typically being the only service that operates 24 hours a day, seven days a week. Nursing staff, that is registered nurses (RNs) and staff who support them providing hands on care, make up the largest element of the health care workforce, accounting for more than half the workforce in many countries (WHO 2016) and is the biggest single cost in running a hospital. For example, in Sweden 71% of the health care workforce are said to be nursing staff (Ridelberg, Roback, and Nilsen 2014).

Determining optimal nurse staffing levels thus epitomises the quality versus costs tensions facing health care providers. Hospital wards need enough nurses on duty to meet patient needs and deliver the care required, safely and to a high standard. Too few staff may lead to care being compromised (Sochalski 2004), work pressures intensified leading to burnout (Aiken et al. 2002), more staff going off sick (Michie and Williams 2003), and costly recruitment and retention challenges (Duffield et al. 2014). Yet providers cannot afford to staff wards with more nurses than they need. With too many nurses, precious resources are potentially wasted, leaving other parts of the hospital, or elsewhere in the system, with too few nurses.

Across Europe, RN staffing levels are not stipulated in law, as they have been in California and parts of Australia (Buchan 2005; Gerdtz and Nelson 2007; Aiken et al. 2010). The number of nurses each ward needs to employ to meet patient needs, and the staffing levels on each shift, is determined locally, by individual unit managers and hospitals. Decision makers need good quality evidence as to the effects, and cost-effectiveness, of nurse staffing in order to make such decisions.

The benefits of a better understanding of the relationships between nurse staffing and patient outcomes are likely to be significant (Griffiths 2012). Needleman and colleagues estimated that in the USA, increasing RN staffing up to the 75th percentile could lead to the potential avoidance of 6,754 in-hospital patient deaths in a year (Needleman et al. 2006). Others have reviewed the economic value of professional nursing in terms of reduced patient complications and the shorter lengths of stay associated with improved nurse staffing levels (Dall et al. 2009).

Conversely, the harm of failing to use evidence to plan nurse staffing can be devastating. Take, for example, the case of Mid-Staffordshire Hospital Trust in England, which hit the headlines in 2009 due to reports of poor standards of care, high levels of patient complaints, low levels of patient satisfaction and higher than expected mortality rates (Healthcare-Commission 2009). The initial reaction of the media was to blame care quality problems on individual nurses. Debate ensued about 'Where had nursing gone wrong?' Were nurses, 'too clever to care'

(Griffiths 2012) or, as phrased ten years earlier, 'too posh to wash' (Hall 2004)? Subsequent inquiries into this hospital, and 14 others that had higher than expected mortality rates, turned attention away from the individual to examine the failings of the systems involved. In a publicly funded National Health Service (NHS), with national professional and organisational regulation, how had patient safety and care quality been so severely compromised, and why had these systemic care failings not been detected earlier? RN staffing levels were identified as a key issue in hospitals with higher than expected mortality rates (Keogh 2013). Sir Robert Francis led a public inquiry into the events at Mid-Staffordshire and the system's failings associated with high mortality rates (Francis 2010). The inquiry found that, along with a host of other problems, staffing levels had played a key part. Striving for financial targets had resulted in the hospital making workforce changes without consideration of the risks to patients (Francis 2010). The inquiry and its recommendations created seismic waves in the health service; the government responded with a raft of policy initiatives aimed at ensuring 'safe' nurse staffing levels in the NHS.

Such scenarios are not limited to England alone. In Sweden, the realisation that the death of a 51 year old woman following heart surgery had not been discovered or reported for several hours, promoted the Swedish Accident Investigation Authority, for the first time ever to examine the system failings of a hospital. They asked: what were the procedural and system errors that enabled such an appalling lapse in care? (Swedish-Accident-Investigation-Authority 2013). The investigation surfaced more general concerns about the degree of consistency in the conditions needed for 'safe care' to be provided in Swedish hospitals. The patient safety challenges related to preventable adverse events in health care, is as much an issue for hospitals in Sweden as for other countries. A review of a sample of 2,000 medical records from the Swedish National Patient Register found that 12.3% had experienced adverse events, 70% of which were categorised as preventable, equating to 8.6% of admissions (Soop et al. 2009). Patient safety legislation was introduced in Sweden in 2011 to reduce adverse events in Swedish health care (Ridelberg, Roback, and Nilsen 2014).

The Macchiarini case at Karolinska University Hospital exposed system failings that jeopardized patient safety and led to the death of two patients, and life threatening complications in a third, following an experimental surgical procedure (Macchiarini-Case-Investigation 2016; Abbott 2016). Many of the factors identified in the investigation resonate with those from the Mid Staffordshire case in England. In a highly competitive context, the safety of patients had been put at risk – guidelines not followed, risks not assessed, and operational safe guards not applied – in the pursuit of a higher level 'prize': status as a 'Foundation Trust' in the case of Mid Staffordshire, and acclaim for surgical innovation and research in the Macchiarini case. In the words of the investigation report: *"Group thinking, bandwagon effects, a very competitive care environment, many informal leaders and deficient knowledge of and respect for rules are some of the factors that may have contributed to the course of events"* (Macchiarini-Case-Investigation 2016) p20.

That there is a relationship between system factors such as RN staffing levels, the quality of care provided, and patient outcomes would come as no surprise to many in health care. Decades of research have sought to elucidate these relationships, looking for the potential effect of RN staffing on patient outcomes. The research has focussed primarily on negative outcomes or 'harms' which could in theory have been avoided: adverse events (medication errors, patient falls), complications (hospital acquired infections, pressure ulcers) and potentially avoidable deaths. 'Avoidable deaths' is a term used to refer to differences in hospital mortality rates that appear to be related to differences in the hospitals and the care provided, as opposed to differences related to patient factors such as diagnosis, co-morbidities, gender, age.

The overall conclusion of research to date is that patients who receive their care in environments with good RN staffing levels are less likely to die from hospital related factors (Kane et al. 2007). Other studies have found that poorer nurse staffing levels are associated with an increased risk that some necessary nursing care is left undone, incomplete or missed (Kalisch, Tschannen, and Lee 2011; Jones, Hamilton, and Murry 2015; Cho et al. 2015).

Whilst the overarching message of the research on RN staffing and patient outcomes is clear, and chimes with the analysis of systems failings put forward in various reports, there remains a lack of clarity in the underlying theory and a lack of detail and specificity about necessary nurse staffing levels that hampers translation into practice. (NICE 2014).

Perhaps some of these 'uncertainties', which are explored in more detail in Chapter 3, contribute to a degree of scepticism expressed by some health managers and policy makers, regarding the evidence on nurse staffing levels and patient outcomes. For example, *"there is no direct correlation between number of staff and good or bad care"* said Harry Cayton, a senior health care regulator of Council for Healthcare Regulatory Excellence in the UK (now called the Professional Standards Authority for Health and Social Care), adding that *"lack of staff is often an excuse for poor care"*(Cayton 2012) p16.

Such views seem at odds with the evidence available, and with experience documented from the high profile instances of health care failings. Yet this perspective – that nurse staffing is an excuse and nothing more – persists. Whilst there may be many different factors influencing the extent to which managers and policy makers know, comprehend and use evidence to inform practice, we cannot rule out the possibility that some of the apparent reluctance may be related to the evidence itself. How complete is the research evidence on nurse staffing levels and patient mortality in hospitals? Can it be relied upon? The first step of this thesis is to look at the research evidence on nurse staffing and patient outcomes, and recent reviews of it, and consider its strengths and weaknesses.

A challenge put forward is that the evidence is partial: much of the evidence comes from North America and so its applicability elsewhere is questionable, it is heavily reliant on cross-sectional studies and rarely takes account of the contribution of other members of the health care team. Potentially, the relationship observed between RN staffing levels and patient

mortality in hospitals could be an artefact related to some unseen and unmeasured 'other' factor. In particular, the omission of medical staffing levels from much research is a significant limitation because medical staffing and nurse staffing can be highly correlated (Griffiths, Jones, and Bottle 2013). In general, to counteract the reliance on cross-sectional data, and strengthen the inference of causality therefore requires evidence of a plausible causal pathway. In the case of RN staffing and patient mortality, we thus need to test if the association that is observed between the two is related to an intermediary factor: e.g. care being left undone. An emerging hypothesis, explored within this thesis, is that when RN staffing levels are low necessary nursing care is less likely to be completed, and that care left undone contributes to an increased risk of patients dying in hospital.

2 AIMS

The overarching aim of this thesis is to investigate the relationships between RN staffing levels, nursing care that is left undone (also termed 'missed care'), and patient risk of death in general acute hospitals.

The focus of the research is on registered nurse (RN) staffing levels. Where the term 'nurse staffing' is used, it may refer to a wide range of nursing workforce characteristics, not limited specifically to RN staffing levels. This includes reference to the full range of nursing staff team members, including nursing assistants and health care support workers, and the relative mix of the team (referred to as the 'skill-mix'). It is also used to refer to a broader range of 'staffing' characteristics beyond 'level'; for example encompassing the education levels of the nursing workforce. A list of abbreviations and key terms is provided at the front of the thesis.

The specific research questions that this thesis aims to address are:

> 1. Are differences in RN staffing levels associated with differences in hospital mortality, even when controlling for staffing levels of doctors and nursing support staff? (Study I)
>
> 2. To what extent is necessary nursing care left undone by RNs on acute wards? (Study II, Study III)
>
> 3. Is there an association between 'care left undone' and quality of care or patient safety environment, as assessed by nurses? (Study II)
>
> 4. Is the prevalence of care being left undone related to differences in RN staffing levels on hospital wards? (Study II & III)
>
> 5. What factors, related to the care context, are associated with variation in the prevalence of 'care left undone'? (Study III)
>
> 6. Do hospitals with higher levels of care left undone have higher levels of hospital related mortality? (Study IV)
>
> 7. Does nursing care left undone by RNs mediate the relationship between RN staffing and mortality in acute hospitals? (Study IV)

It is hypothesised here that care left undone may occur as a consequence of lower nurse staffing levels and may in turn lead to higher patient mortality: that is to say care left undone may mediate the relationship between RN staffing and mortality.

Across the literature on nurse staffing and patient outcomes (described in the Background), few studies refer specifically to the exact levels of RN staffing being observed; findings are reported in terms of the strength of the observed relationship without referring to the exact levels or

'dose' of nursing. Yet this is necessary knowledge if research findings are to be useful to health service managers, senior nurses and policy makers who are making decisions about nurse staffing levels required. Thus an additional objective of this dissertation relates to how the findings are reported. Wherever possible, the specific RN staffing levels (or 'dose') being examined in addressing the questions above will be made explicit.

Whilst the theories, research and emerging findings may have implications for other settings, the aim of this thesis is to focus on the context, care and outcomes of the adult in-patient population of general acute hospitals.

3 BACKGROUND

This chapter provides an overview of current knowledge regarding the relationships between nurse staffing levels, mortality rates, and nursing care left undone. I start by looking back at the evolving interest in the relationship between nurse staffing and patient outcomes, and identify two sides to the same research coin: studies that seek to explain differences in mortality rates that consider nursing as a factor, and studies that seek to identify the impact of nursing, that examine patient outcomes such as mortality rates. The intention is to make explicit the evolutionary path of the research, the questions it has sought to address and the approaches deployed.

I then outline some of the research that has explored nurse staffing in relation to two key themes of the dissertation: hospital mortality (as a patient outcome) and care left undone (as an outcome of staffing, and potential mediator between staffing and mortality). In addition to drawing on numerous significant reviews that summarise and appraise published research, I refer to specific studies in order to explore some of the methodological challenges in this field, how they have been overcome, and highlight what is known and, equally important, some of the uncertainties that remain.

The chapter closes with a reflection on the theories put forward to make sense of the relationships observed between nurse staffing and patient outcomes such as mortality.

3.1 NURSING INPUTS – PATIENT OUTCOMES: HISTORICAL PERSPECTIVE

The issues of nurse staffing levels and the relationship with care quality and patient mortality are not new challenges for hospital managers, nor is this a new topic for investigation.

In 1930 the Lancet published a letter from Dr Esther Carling, the medical superintendent of a sanatorium, raising the alarm about the lack of nursing staff. This letter prompted the Lancet to launch a commission into the causes of the nursing shortage (Lancet-Commission 1932). Earlier still, the roots of modern day nursing design are grounded in Florence Nightingale's analysis of causes of death during the Crimean War, revealing how much mortality was related to the conditions of the hospitals, as opposed to combat (Nightingale 1858). A hundred years later, the conditions needed for safe and effective care in acute hospitals, including the nurse staffing required, was a topic of interest in the USA (Levine 1961). Differences in hospital mortality rates, and an examination of nurse staffing have continued to be studied in the fifty years since.

Several studies followed Levine and colleagues' work, looking at variation in hospital mortality rates, in relation to variation in other hospital characteristics. Nursing was included alongside a number of other features. For example Hartz et al looked at a range of characteristics in relation to case-mix adjusted mortality rates across 3,100 hospitals (Hartz et al. 1989). They examined occupancy rates, board certification of physicians, pay-roll expenditure per bed, and

nursing skill-mix – that is, the proportion (typically expressed as a percentage) of nursing staff that are 'registered' and have at least 3 years of full-time nursing education. Contrasting the upper and lower quartiles according to the percentage of nursing staff that were RNs, they reported a significant difference (P<0.0001) in hospital mortality rates (113 vs 119 per 1000).

In contrast to these hospital outcomes studies that touched on nursing, many studies in the 1980s gave an in-depth view of the nursing input, and particular approaches to organising nursing care. Relatively few of these however included objective measures of patient outcome, such as mortality rates (Buchan and Ball 1991). Many of these studies looked at systems of organising nursing – such as primary nursing where each nurse has a set of patients that they have 'primary' responsibility for, or nursing led care, or all RN-units (Sellick, Russell, and Beckmann 1983; Pearson 1988; Bowers 1989). Whilst such systems clearly have implications for how units are staffed, the primary focus of the research is on the system of care overall. The specific contribution that differences in nurse staffing levels may make, is thus difficult to isolate from the bundle of characteristics that constitute the model of care delivery being examined; they rarely, if ever, report comparative staffing levels.

In the early 1980s many hospitals in the USA were once again facing severe recruitment and retention challenges. According to Whelan, nursing shortages are a recurring problem in health care (Whelan 2000). The context of nursing shortages prompted interest in why some hospitals appeared better able than others to both attract and retain nurses. The 'American Academy of Nursing's (AAN) Task Force on Nursing Practice in Hospitals' studied 163 hospitals to identify and describe the characteristics of hospitals with reputations for being good places to practice nursing. What was it that made these hospitals 'magnets' to nursing staff? Research identified that one of the characteristics of 'Magnet' hospitals was nurses' satisfaction with the adequacy of resources including sufficient staff on duty to provide the care needed (Kramer and Schmalenberg 1988).

Up until the early 1990s, research themes on the organisation of hospitals – including nurse staffing – and on variation in mortality rates had *"developed independently of one another for the most part"* (Aiken, Smith, and Lake 1994). In a review of the costs and benefits of nursing, published in 1991, a colleague and I concluded that *"there is no body of literature on the costs and benefits of nursing, but rather a series of often incompatible parts"* (Buchan and Ball 1991) p34. At that time, whilst much research purported to address the 'value of nursing care', we found few studies that examined nursing inputs in relation to measures of patient outcome, whilst also considering costs. The evidence base was fragmented, and dominated by research from North America, with very little research based in Europe.

In 1994, Aiken and colleagues revisited the Magnet hospital research and published a paper that started to address the question: are hospitals that are good for nursing staff also better for patients? (Aiken, Smith, and Lake 1994). Their analyses found that 'Magnet' hospitals had significantly lower case-mix adjusted mortality rates, than a sample of matched comparator

non-magnet hospitals. The study marked a step-change in the research in this area, looking explicitly at hospitals in terms of their nursing characteristics and at an empirical measure of patient outcomes: case-mix adjusted mortality.

A few years later in 1999, the NIH-funded 'International Hospital Outcomes Study' (IHOS) expanded and developed on this work, replicating a research design in five countries: Canada, England, Germany, Scotland, and the USA (Aiken et al. 2001). The study sought to examine a range of organisational characteristics that may contribute to positive nursing and patient outcomes, looking at characteristics of the environment in which care was delivered, and examining issues such as nurses' autonomy, leadership, and teamwork, in addition to nurse staffing factors. The results from England on the relationship between nurse staffing and mortality rates, and other outcomes such as job dissatisfaction, burnout, and failure to rescue, were published in 2007 (Rafferty et al. 2007). The results – that hospitals with RN staffing in the top 25% had significantly lower risk of adverse outcomes compared to those in the bottom 25% – had considerable impact, being some of the first to present findings from outside of North America on the relationship between nurse staffing and patient mortality (Rafferty et al. 2007).

In the wake of the 1999 IHOS, a proliferation of studies adopting a similar design followed, looking at nurse staffing in different settings and in different countries (Needleman et al. 2002; Sasichay-Akkadechanunt, Scalzi, and Jawad 2003; McGillis-Hall, Doran, and Pink 2004; Estabrooks et al. 2005; O'Brien–Pallas et al. 2010; Kutney-Lee, Sloane, and Aiken 2013; You et al. 2013).

What these studies have in common with one another is that they have used a retrospective observational design. Some used administrative data to provide a measure of Nursing Hours per Patient Day (NHPPD). Many have undertaken large-scale cross-sectional surveys of nurses to capture nurse reports of staffing at the ward/shift level and other workforce attributes, nurse views of their practice environment, nurse burnout, and intention to leave, and nurse reported quality of care. Most include measures of hospital characteristics such as size, presence of an ITU, and teaching hospital status. Typically multivariate regression analyses are used to examine nurse staffing in relation to patient outcomes derived from administrative datasets, with differences related to patient risk factors taken into account within the model.

A few follow the original template based on Aiken and colleagues' design, combining an examination of hospital and nursing characteristics with administrative data on patient mortality, adjusting for patient characteristics in order to explore the relationship between aspects of nurse staffing and hospital related mortality.

McHugh and colleagues revisited earlier work looking at mortality rates in Magnet hospitals in order to see how much of the effect was a result of better nursing characteristics (McHugh et al. 2013). They reported that the lower mortality rate found in Magnet hospitals was indeed

largely, but not wholly, attributable to nursing characteristics, such as nurse staffing levels and practice environment. There was, nonetheless, residual benefit to Magnet hospitals in terms of lower risk of hospital related mortality, beyond the nursing characteristics measured (McHugh et al. 2013).

The RN4Cast study, which is the source study for this dissertation, follows in this tradition but on a larger scale, covering 12 countries (Sermeus et al. 2011).

Some papers published from these studies have focussed on exploring the relationships between variables captured in the nurse surveys, without connecting the data to empirically derived measures of patient outcomes. For example, some studies have examined staffing variables and practice environment characteristics in relation to nurse reported outcomes, such as job satisfaction, intention to leave, burnout or care quality (Sheward et al. 2005; Heinen et al. 2013; Zúñiga et al. 2015; Laschinger and Leiter 2006). These findings have informed the development of theories regarding how attributes of nursing organisation may be related to one another and to different nurse reported outcomes, as discussed at the end of this chapter (Tourangeau 2005).

These studies have established that there are numerous significant associations between the different nursing characteristics and outcomes. The evidence base has established, with a degree of confidence, the existence of multiple associations between RN staffing and a range of variables. However, by simply drawing a line between two dots, we have little sense of how all the dots join: we have yet to join and number the dots, to reveal the entire picture. The somewhat piecemeal examination of one set of variables in relation to another does not offer a more consolidated view of the relative importance of different nursing characteristics, nor does it reveal how different variables may be interacting, moderating, or mediating the relationships described. Are low nurse staffing levels associated with poorer patient outcomes due to a reduction in nursing capacity to deliver care, with increased workload pressure causing higher levels of burnout and sickness absence, as Tourangeau suggests (Tourangeau 2005), or is it that a reduction in the presence of nurses results in less surveillance and monitoring which results in patient deterioration going undetected, as Aiken and proponents of a 'failure to rescue' theory described (Aiken, Clarke, and Sloane 2002)? Whilst researchers have put forward both of these as possible explanations for understanding the observed relationship between RN staffing and patient outcomes, they have rarely directly tested the hypotheses on offer, or how they may interact.

A recurring limitation reported across many of the studies is the cross-sectional design (Shekelle 2013b). The inference of causality requires a strong theoretical explanation as to why the observed association is likely to be in a particular direction and likely to be causal, and not arising in response to an unseen 'other' variable. Aside from a plausible theoretical explanation, inference of causality is supported by: minimising risk of confounding, temporal precedence, and evidence of a dose-response effect (Pearl and Verma 1995; Holland 1986). In relation to

minimizing the risk of confounding, little of the published research relating RN staffing to mortality has considered the input of other staff groups, such as medical staff and support workers.

The evidence base is not, however, entirely limited to cross-sectional studies which use surveys of nurses as a source of staffing data, and examine the association between measures of nurse staffing and measures of patient outcome, across a defined period of time. A few studies have been published that use different designs. Their differences make them critically important to the evidence base.

Research in New South Wales, Australia adopted a longitudinal design using NHPPD to measure staffing levels (on medical and surgical wards) with a cross-sectional study to capture data on nursing work environments (Duffield et al. 2011). In this the research 11 patient outcomes were examined including: patient falls, medical errors, urinary tract infections, and other adverse events plus 'failure to rescue' – defined as death following sepsis, pneumonia, GI bleeding or shock, following the method previously adopted by Silber and colleagues (Silber et al. 1992). Higher levels of RN staffing were associated with lower levels of adverse events. The differences persisted for three outcomes, when examined through a multivariate analysis of matched units: central nervous system derangement, and urinary tract infection and failure to rescue.

Shekelle identified the work of Needleman and colleagues published in 2011 as a key study (Needleman et al. 2011; Shekelle 2013b). Whilst it is also an observational study, its strength lies in its retrospective design and use of survival analysis (using Cox proportional-hazards regression models) which allowed temporal precedence to be established: the outcomes observed follow the inputs observed. Nurse staffing was measured through routinely collected administrative data, and was thus recorded for every shift, covering 176,696 eight-hour shifts from 43 units in one hospital. Low RN staffing was defined as shifts in which the number of nursing hours per patient day was eight or more hours less than the planned level. A significantly increased risk of mortality was observed *after* periods of exposure to low staffing. An additional merit of the study is that it makes explicit an objective measure of 'low staffing' rather than relying on a relative measure of RN staffing related to an arbitrarily defined reference point. Risk of patient death is increased after exposure to an empiric RN staffing deficit: shifts where the number of nursing hours per patient day (NHPPD) is at least eight hours fewer than estimated as required – i.e. a shortfall of one RN.

The study has been widely cited by academics, clinicians, and national bodies. For example, in a review undertaken by the Royal College of Surgeons of England examining the care context of 'forgotten patients' (that is technically 'low risk' surgical patients who unexpectedly deteriorate), they refer to the Needleman et al study measure of '8 hours less RN care than expected' as a benchmark in reviewing the post-surgical care of patients (RCSE 2011).

Several key features distinguish the design of the study undertaken by Needleman and colleagues. They used repeated measurements, the unit of measurement and analysis is at ward level not at hospital level, and the nurse staffing levels are captured as nursing hours per patient day from administrative rostering data rather than relying on nurse reports.

3.2   MORTALITY IN HOSPITALS

In the literature seeking to examine the impact of nursing on patients, a variety of measures or indicators have been adopted. Mortality rates are frequently used as a measure of patient outcome.

3.2.1   Hospital mortality rates as an indicator

Hospital mortality rates measure the number of deaths within a hospital. The cause of patient death in a hospital can be attributed to two types of factors – patient related and hospital care related. The majority of patients who die in hospitals do so as a result of a combination of their presenting condition (which brought them to hospital), underlying health factors ('co-morbidities'), and their age. By 'case-mix adjusting' mortality rates for hospitals, data-analysts seek to take account of patient related factors in order to identify mortality variation that relates to differences in the hospital care provided, not simply to differences in the mix of patients present. This is the essence of the idea of 'avoidable mortality': deaths that relate to factors other than patient factors, that are related to the context and care provided within the hospital, which may potentially have been avoided (Kossarova et al. 2009; Mackenbach, Bouvier-Colle, and Jougla 1990; Nolte and McKee 2004; Vincent et al. 2004; Rutstein et al. 1976).

Hospital mortality rates have been used as a metric for quality improvement. Birkmeyer et al make the point that directly measuring 'outcomes' such as mortality rates – as opposed to process or structural variables, applying the Donabedian paradigm of quality – has the advantage that it has direct and obvious relevance; it addresses the 'bottom line' of surgery performance (Birkmeyer, Dimick, and Birkmeyer 2004).

However, there are some concerns that routinely collected administrative data may not be sufficiently well-coded to provide an accurate measure of hospital related in-patient death (McKee, Dixon, and Chenet 1994). In comparing risk-adjusted outcomes from routinely collected data to that derived from clinical databases, Aylin and colleagues reported in 2007 that this approach could be used to predict risk with a similar level of differentiation as clinical databases. They concluded that the creative use of case-mix adjusting administrative may offer a useful approach to monitoring health care performance (Aylin, Bottle, and Majeed 2007). One such approach has been to focus on mortality rates for a particular set of patients rather than across a whole hospital: for example by examining mortality for patients who have low inherent risk of death, or for a group who are undergoing similar procedures. The objective of narrowing the focus in this way, is to reduce the amount of variation between observed and

expected mortality rates that is likely to be related to patient differences, and increase the likelihood that differences see are attributable to service level factors.

Thus whilst the relative rareness of patient death following surgery restricts the usefulness of mortality rates as an indicator of an individual surgeon's performance, by grouping procedures together it can be used as an indicator of hospital performance (Birkmeyer, Dimick, and Birkmeyer 2004).

When Hogan and colleagues retrospectively reviewed the case records of 1000 adults who died in 10 acute hospitals in England in 2009, they found that the incidence of what might be regarded as 'preventable' or 'avoidable' hospital deaths was about 50% of the average 5% mortality rate. They concluded that *"a focus on deaths may not be the most efficient approach to identify opportunities for improvement given the low proportion of deaths due to problems with healthcare"* p737 (Hogan et al. 2012).

But note that this reservation is applied to the use of case-mix adjusted mortality as an indicator for *quality improvement* purposes within local systems. The choice of measures in a research context has potentially different requirements; there is no need to make assessments that translate to the individual surgeon or even the individual hospital. Thus there is greater opportunity for aggregation and examining mortality across groups of hospitals – as done in the early work comparing Magnet hospitals to non-Magnet hospitals, for example (Scott, Sochalski, and Aiken 1999).

There is considerable variation in rates of case-mix adjusted mortality between hospitals (Ghaferi, Birkmeyer, and Dimick 2009; Tourangeau 2010). The key to unlocking mortality rates as an outcome measure for research purposes, is to take account of patient related factors, leaving behind the non-patient, hospital related factors. Methods for 'case-mix' adjustment have evolved beyond including a measure of the mix of cases. Analytical techniques have been refined to take risk factors into account at the level of the individual patient. Risk of death is estimated directly, based on the age, gender, comorbidities etc. of each patient, and account for 'clustering' (Silber et al. 1992; Silber and Rosenbaum 1997; Aylin, Bottle, and Majeed 2007; Cohen et al. 2013; Goranitis and Fischer 2015).

3.2.2 Nurse staffing and hospital mortality: evidence reviews

Lankshear and colleagues undertook a systematic review of published research and a policy analysis (which included interviews) to examine studies of the relationship between the nursing workforce and case-mix adjusted patient outcomes in acute hospitals (Lankshear, Sheldon, and Maynard 2005). They reported that higher nurse staffing levels and a richer skill-mix (i.e. larger proportion of RNs in the nursing workforce) were associated with better patient outcomes, although they noted that the effect-size could not be estimated reliably.

In 2007, Kane and colleagues systematically reviewed the evidence on RN staffing and adverse events in hospitals. They pooled data from 28 studies, examining the risk of in-patient mortality in acute hospitals in relation to RN staffing levels, in order to undertake a meta-analysis. Their conclusion, in common with Lankshear et al, was that increased RN staffing levels were associated with lower odds of hospital related mortality and adverse patient events (Kane et al. 2007).

In 2010 Tourangeau reviewed research that had used mortality as an indicator in relation to nursing structures or nursing processes of care (Tourangeau 2010). The review was not restricted to any particularly setting, nor to an examination of nurse staffing, yet all 17 of the papers identified related to acute hospital care, and all included a measure of nurse staffing. Fourteen of the studies reviewed reported a statistically significant association between nurse staffing and mortality rates. Whilst a strong relationship was found, Tourangeau concluded that given the reliance on cross-sectional studies, more research was needed to examine theoretical explanations of the relationship between higher nurse staffing and lower hospital mortality, and test if this relationship may be causal.

Shekelle's review in 2103 built on the preceding reviews by Kane et al and Tourangeau (Kane et al. 2007; Tourangeau 2010). Like Tourangeau, he noted the reliance on cross-sectional data, and highlighted the risk of confounding: that the better outcomes observed may be due to hospitals being "better" in other ways, in addition to their RN staffing. He identified 15 new studies that related patient ratios and patient ill-health or mortality. He concluded that: *"The strongest evidence supporting a causal relationship between higher nurse staffing levels and decreased inpatient mortality comes from a longitudinal study in a single hospital* (Needleman et al. 2011*) that carefully accounted for nurse staffing and patient comorbid conditions and a meta-analysis* (Kane et al. 2007) *that found a "dose–response relationship"* in observational studies of nurse staffing and death*"* p404 (Shekelle 2013b).

In 2013 in England, following inquiries into unexpectedly high death rates at some acute hospitals, the Department of Health tasked the National Institute for Health and Care Excellence (NICE, an independent body) with the development of guidance on 'safe' nurse staffing on acute hospital wards (Nursing-Times 2013). To support the development of guidelines, and following the principles of evidence based practice, NICE commissioned a team of researchers (which included me) to review the evidence on nurse staffing and patient outcomes (Griffiths et al. 2014).

Our review built on the strategy for Kane's (2007) systematic review of nurse staffing and outcomes. The scope of the literature identified as relevant was vast. This allowed us to apply strict inclusion criteria and limit the papers for review to those that had most direct relevance. The papers reviewed were thus limited to: general acute medical and surgical settings (including older peoples' wards), from 1993 onwards, studies that controlled for the contribution of other members of the nursing team (including nursing support staff), and studies

that measured nurse staffing at the ward or unit level (not hospital). The outcomes included were:

a. Mortality, and 'Failure to Rescue' (i.e. death following complications)

b. Avoidable harms (such as hospital acquired infections, serious preventable safety incidents, pressure ulcers, falls, urinary tract infections, venous thromboembolisms)

c. Care omissions and errors: observing/monitoring patients, fundamental care done/undone, drug administration errors

d. Patient reports: complaints, patient satisfaction, patient reported outcome or experience

Thirty-five primary studies met the stringent inclusion criteria for full review. Of these, nine studies looked specifically at the relationship between RN staffing levels and mortality. In four of the studies, a statistically significant association between lower nurse staffing and higher mortality rates was reported (Needleman et al. 2011; Blegen et al. 2011; Sales et al. 2008; Sochalski et al. 2008). Whilst the results from the other studies were not statistically significant (e.g.Kutney-Lee, Sloane, and Aiken 2013), it is nonetheless worth noting that none of the studies found a statistically significant relationship in the opposite direction. From these findings we reached the same conclusions as Kane and colleagues: there is evidence of a clear association between nurse staffing levels and mortality in acute hospitals (Kane et al. 2007; Griffiths et al. 2014).

Whilst that overall conclusion regarding the evidence base was supported, the review and quality appraisal process that we applied unearthed a number of design weaknesses in the research in this field. Each study was critically appraised in terms of its internal validity, external validity, and risk of bias, using an adapted version of the NICE quality appraisal checklist, designed for use with quantitative studies where correlations are reported. A summary of the appraisal is presented in Table A.1 in the Appendix. Only four of the 35 papers reviewed were assessed as relatively strong for both internal and external validity.

Ultimately the evidence reviewed for NICE, did not lend itself to the purpose for which it had been reviewed: to develop safe staffing guidelines for general acute hospital wards. Lack of specificity regarding the nurse staffing levels examined, and limited research relating to the UK or EU context limited the extent that the evidence could be used to provide concrete 'safe staffing' guidelines. We described the strengths and weaknesses of the evidence base, and discussed potential sources of bias in the estimation of causal effects in studies relating nurse staffing to patient outcomes (Griffiths et al. 2016). Three issues were highlighted in the studies reviewed:

- Omitted variables: few studies take account of the staffing levels of doctors; some studies do not consider differences in the mix of patients.

- Simultaneity: other factors may simultaneously influence staffing and outcomes.

- Common-method variance: a reported association could be a result of both things being measured using the same method – such as a questionnaire. For example, nurses who are negative about staffing may generally be negative, and so are also negative about quality of care.

3.3 CARE LEFT UNDONE

Much of the research on nurse staffing and mortality described above has focussed on establishing that there is a significant association between RN staffing and case-mix adjusted mortality, which remains even when controlling for other factors. Explanatory theories put forward have suggested that lower staffing levels may lead to less surveillance, with potentially fatal deterioration going undetected, or that when there are lower levels of RN staffing, the capacity of the nursing team to provide care is reduced. Both of these theories potentially connect to a single underlying construct: necessary nursing care being left undone at times of insufficient RN staffing – whether that care be vital signs observations, surveillance or other fundamental care (Kitson et al. 2010; Feo and Kitson 2016). Hypothetically, care left undone may be both a consequence of lower RN staffing and precursor to adverse patient outcomes, including mortality.

3.3.1 Measuring 'Care left undone'

'Care left undone' is also described in the literature as 'missed care', 'incomplete care', 'unfinished care' and 'implicitly rationed care'. Jones and colleagues, who comprehensively reviewed the research on care left undone, make the point that *"few care processes reach patients without first passing through the hands of nurses"* p1122 (Jones, Hamilton, and Murry 2015). If the 'flow' of nursing care is blocked, there is an increased risk that patients may not receive all the care planned, with some necessary care being unfinished or left undone. Incomplete or missed care represents a form of health care 'under-use' which, argues Reason, is the most common cause of quality problems in health care, more so than 'over-use' or 'misuse' combined (Reason 1998). Yet despite a growing body of literature within nursing over the past decade, missed nursing care as a form of 'under-use' has received little attention within the patient safety world. Jones and colleagues suggest that the lack of awareness may, in part, be due to the inconsistencies in terminology (Jones, Hamilton, and Murry 2015).

Three main instruments have been used to measure care left undone, rationed care or missed care. The 'Care Left Undone' scale (also referred to as 'Tasks Undone') is used by the RN4Cast study, and is based on a scale first used in the International Hospital Outcomes Study (Aiken et al. 2001). A second measure, that is closely related, stemming in part form that early work by Aiken and colleagues, is the Basal Extent of Rationing of Nursing Care Instrument BERNCA (Schubert et al. 2007). A third measure, the 'MISSCARE survey', was developed by Beatrice Kalisch and colleagues (Kalisch and Williams 2009) drawing on qualitative research on the activities and reasons for care being missed (Kalisch 2006).

Whilst all three include an inventory of nursing care activities and capture responses through self-report surveys, they each approach the measurement of care undone slightly differently: variation in recall period, response scale (frequency not completed versus a dichotomous yes or no), the number and range of activities, and the inclusion of 'reason' for care undone. A limitation of all three instruments (and derivatives of them) is the reliance on self-report; there has been no evaluation to establish congruence between reported missed care and directly observed missed care.

MISSCARE is the only instrument that has been adapted to enable its use by patients, to capture their perceptions of care that is unfinished (Kalisch, McLaughlin, and Dabney 2012); the others are used by nursing staff or RNs only.

3.3.2 Causes of care left undone: lack of time

Time scarcity is frequently identified as the primary cause for care being left undone or missed. Time scarcity itself is a reflection of workload, which in the case of ward nurses, is determined largely by whether staffing levels are sufficient relative to both the number and dependency of patients that are cared for. Nurse reported care intensity was found to be predictive of unfinished care, but none of the studies reviewed by Jones examined had assessed nursing intensity formally using a recognised system (Jones, Hamilton, and Murry 2015). In studies using the MISSCARE Survey, 'inadequate labour resources' (particularly unexpected workload increases due to changes in the number or mix of patients) was the most significant factor identified as contributing to missed care.

3.3.3 Consequences of care left undone

The majority of research related to care left undone has examined the causes or antecedents of missed care; few have focussed on the possible consequences of missed care. Of the 54 studies that Jones and colleagues reviewed on missed care (Study II from this thesis included), only two examined patient outcomes (Jones, Hamilton, and Murry 2015). One of these looked at patient falls (Kalisch, Tschannen, and Lee 2012), whilst the other examined rationed care and hospital mortality, using data from eight Swiss acute hospitals (Schubert et al. 2012). Schubert and colleagues concluded that their small sample size and study design limited the generalizability of their findings and did not enable them to test the influence of care left undone on the observed relationship between staffing and mortality. The challenge of small hospital sample sizes is noted by others in relation to using hospital mortality as an outcome measure (Dimick, Welch, and Birkmeyer 2004).

The impact of missed care, through the eyes of patients and relatives who have reported adverse events to Swedish 'Medical Responsibility Board' (Hälso-och sjukvårdens ansvarsnämnd: HSAN), was highlighted by Andersson and colleagues (Andersson et al. 2015). Through qualitative analysis of 242 adverse events in nursing, they found that adverse events were caused not just by errors made by health care staff, but through care that had been omitted.

Increased attention to the implications of missed care is needed, they concluded, in addition to an investigation of the relationship between missed care and lack of staff with relevant competence.

Bruyneel and colleagues found that hospitals with lower levels of clinical care left undone (a subscale based on aggregating seven activities of clinical care left undone) had higher levels of patient satisfaction, and that clinical care left undone mediated the relationship between nurse staffing and patient satisfaction (Bruyneel et al. 2015). This is one of the first to have tested the mediating effect of care left undone between RN staffing and a patient measure.

3.4 THEORETICAL MODELS OF NURSE STAFFING AND PATIENT OUTCOMES

In 1995, Mark described the research examining patient outcomes in relation to nursing as being atheoretical; a 'black-box' approach to science (Mark 1995). She observed that many studies report on the inputs and outcomes, and correlations between the two, without putting forward and testing a theoretical model to explain the mechanisms through which these relationships may be operating (Mark 1995). Ten years later, despite a growing body of research, Mark and colleagues still saw the absence of theory as a challenge:

"There is considerable work to be done in explicating the theory underlying empirical studies of the relationship between nurse staffing and outcomes. Key constructs must be placed within a theoretical context, proposed causal mechanisms underlying the empirical or hypothesized relationships must be identified, and critical mediating and moderating variables must be recognized."(Mark, Hughes, and Jones 2004) p11.

Shekelle (2013) outlined four conceptual frameworks put forward to understand the relationships between nurse staffing and in-patient mortality (Shekelle 2013a), proposed by Aiken and colleagues, Tourangeau, Thornlow and colleagues, and Despins and colleagues.

The model put forward by Aiken and colleagues in 2002 proposes that nurse staffing levels and skill-mix impact on mortality indirectly through two routes (Aiken et al. 2002). Firstly through the association between nurse staffing levels and broader organisational support for nursing, which is seen as leading to better nurse outcomes. Broader organisational support (including staffing levels) is thus put forward as a predictor of better patient outcomes. Secondly, nurse staffing is linked to outcomes through improved nursing surveillance, via the 'care process'.

In 2005 Tourangeau worked on developing a theoretical model of the determinants of mortality, informed by research findings from multiple studies (Tourangeau 2005). In this model, nurse staffing is one of five characteristics that are hypothesised to indirectly affect patient mortality. Tourangeau suggests that the total nurse staffing 'dose' (all nursing staff, not just RNs) impacts on nurse burnout, which relates to nurse satisfaction, which influences nurses 'capacity to work' (measured by scheduled shifts that are missed due to sickness), and that it is through 'capacity to work' that the nursing 'dose' influences mortality. Alongside the effect of

the 'dose' of nursing staff, she hypothesises a separate direct relationship between nursing skill-mix and patient mortality, although no mediators are put forward for this relationship. The paper provides a detailed account of the many factors identified from research as having an association with patient mortality, and puts forward a complex model of indirect and direct relationships. Yet the hypothesis relating specifically to how RN staffing levels might impact on in-patient mortality rates is unclear. Two distinct pathways are put forward: the 'dose' is considered in terms of the total volume of all nursing staff (including nursing support staff), and a separate skill-mix variable is used to estimate RN input. It is not apparent, in the theoretical model put forward, how RN staffing (i.e. the RN specific 'dose') may impact on mortality, and what factors might mediate or moderate that relationship.

Thornlow and colleagues proposed a model that applies specifically to post-operative outcomes (Thornlow, Anderson, and Oddone 2009). They summarise a range of factors that are likely to contribute to the risk of post-operative respiratory failure (including surgical procedure, type of anaesthesia etc.). The model includes staffing as one such factor, and interventions such as nursing surveillance as another related factor. This idea, that nurse staffing may impact on patient mortality through levels of surveillance that can be provided, is a common thread across the literature. It is found informally in discussions of research findings to offer explanations for the observed relationship. The hypothesis is this: the presence (or absence) of sufficient numbers of RNs determines the level of surveillance that is possible, and this surveillance is required to detect and respond to changes in a patient's condition, and to avoid deterioration that could prove to be fatal. Surveillance as a mechanism is encapsulated in the notion (and subsequent measurement of) 'failure to rescue' (Silber et al. 1992; Griffiths, Jones, and Bottle 2013).

Despins and colleagues produced a model which focuses specifically on nurses' ability to detect and respond to changes in patients' conditions that put patients at risk (Despins, Scott-Cawiezell, and Rouder 2010). In this model, nurses' sensitivity and ability to respond to signs of deterioration are seen as being determined by the broader organisational context that shapes attitudes to failure, reluctance to simplify, and operational sensitivity. The model does not include how RN staffing levels may impact on nurses' surveillance and response function.

Since Shekelle's summary of conceptual models relating nurse staffing to patient outcomes was published (Shekelle 2013a), Bruyneel and colleagues have put forward an alternative model. They have collated the 'isolated findings' from previous research examining different aspects of the relationship between nursing inputs and patient outcomes and put forward an explanatory framework that builds on Donabedian's (Donabedian 1988) determinants of quality: structure, process and outcomes (Bruyneel et al. 2015). Clinical care left undone is put forward as a mediator between RN staffing and patient satisfaction with care, with level of nurse education exerting a moderating influence between RN staffing and clinical care left undone. Whilst this study looked at a different outcome – patient satisfaction rather than

mortality – it one of the only studies to have examined care left undone as a mediating factor in the relationship between RN staffing and outcomes.

Despite the large number of studies establishing associations between nurse staffing and patient outcomes, each putting forward possible interpretations for the relationships observed, these interpretations have rarely been stringently formulated. Models of mediation or moderation to support theories regarding the relationship between nurse staffing and patient mortality have not been fully tested.

4 DEVELOPMENT OF A THESIS

As Chapter 3 demonstrates, at the outset of this thesis, the existence of an association between RN staffing level and risk of hospital related patient mortality had been established (Kane et al. 2007; Shekelle 2013a). Despite the large number of studies undertaken however, the evidence base is summed up as *"voluminous but to a great extent stagnant literature"* (Sales 2015) p241. New directions were needed.

In developing a proposal for my doctoral research, two areas of inquiry helped to illuminate potentially meaningful research questions for my upcoming research; that is questions that had the potential to enhance our knowledge of the relationships between nurse staffing and patient outcomes, and those that might identify possible mechanisms for the observed relationship.

Firstly, a review of the evidence base and identification of knowledge gaps revealed that very little research had examined the relationship between RN staffing and hospital mortality whilst taking account of the contribution of other care givers: doctors and nursing support staff. Secondly, findings from previous research studies had only limited applicability to policy and practice. The existence of a relationship had been repeatedly demonstrated, but the nature of that relationship, and underlying mechanisms that might explain it, had not been made explicit.

My initial focus for my doctoral study was to be on discerning the shape of the relationship, to see if there may be a 'tipping point' or 'threshold' between the level of RN staffing associated with lower risk care compared to higher risk care. The aim of my research proposal at the outset was thus to explore the relationship between RN staffing levels, care provided, and patient mortality to discern if such a staffing 'threshold' existed, and if so whether this varied in different contexts, and different countries, taking Sweden and England as examples. Based on previous research, 'care left undone' (or 'missed care') looked likely to offer a relatively sensitive measure of staffing adequacy, compared with more distal measures of outcomes such as mortality, in that it was frequently observed, and level of missed care varied according to nurse staffing characteristics.

To describe the territory of the research, midway through my doctoral studies I put forward a diagram to illustrate the putative relationships among the variables being examined (Figure 1).

Figure 1: Nurse staffing levels, care (un)done, and patient outcomes – a model
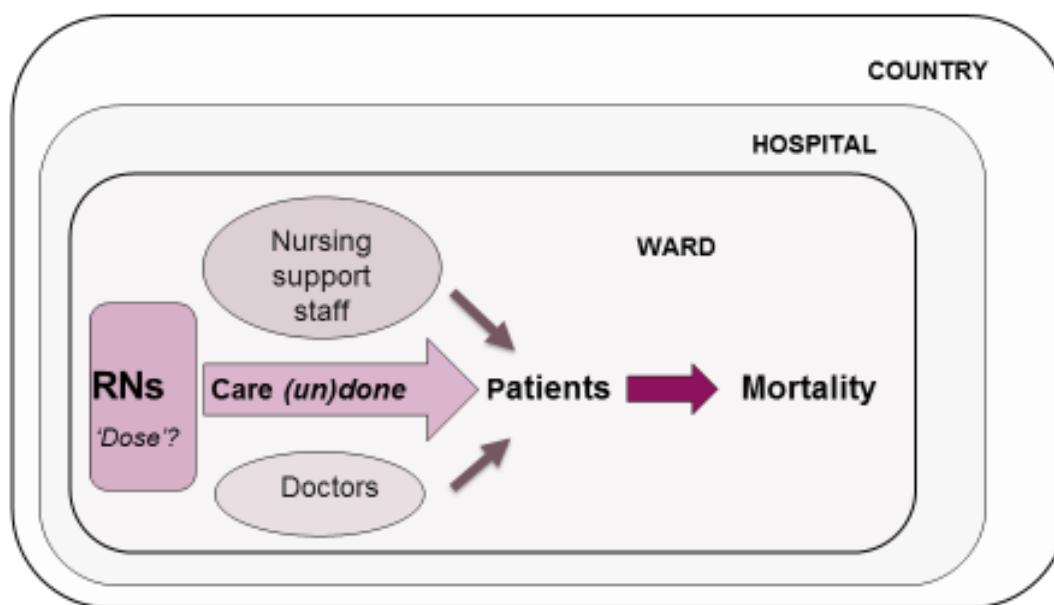
In the diagrammatic representation I shared, 'care done/undone' was placed between RN staffing levels on one side, and patient outcomes, such as mortality, on the other. To illustrate the contributions of other staff to patient care, separate elements were added: one for the staffing level of doctors and another for the staffing levels of nursing support workers. To represent the practice environment in which staffing, care provision and patient outcomes occur, these relationships were placed within a box representing the ward/hospital context, and that context was placed in a bigger frame, which represented the country care delivery context.

The role of 'care left undone' as a consequence of lower staffing, and as a possible precursor to poor patient outcomes, was implied by the diagram. To pursue the original goal of the research, the first three studies focussed on:

- Identifying if a relationship between RN staffing remained when the staffing levels of doctors and support workers were included in multi-level multivariate analyses (Study I)

- Making explicit empirical nurse staffing levels – expressed using metrics that were meaningful in practice – associated with risk of care left undone (Study II, Study III), and with risk of hospital related mortality (Study I)

- Examining 'other factors' that influence each of the two variables: nurse staffing and care left undone, and the relationship between the two, and contrasting the findings from two countries to reveal differences and similarities in the relationship between nurse staffing and care left undone in different contexts (Study III).

During the course of the research and in response to discussion at my doctoral half-time review, it became apparent that this 'diagram', which I had put forward simply as an aid to outline the

research and the variables of interest, depicted a hypothetical causal pathway between RN staffing and outcomes.

Underlying my research was an untested hypothesis: that care left undone by RNs mediates the relationship between RN staffing levels and hospital related patient mortality. This hypothesis presented a viable theoretical connection between nurse staffing levels and patient outcomes. The final element of my thesis was thus redirected to test this hypothesised causal pathway, and determine if care left undone predicts hospital related mortality, and mediates the relationship between RN staffing and mortality.

Realisation of the importance of examining care left undone – not just as an indicator of RN staffing level adequacy, but as a potential predictor of hospital related mortality – thus shaped the second half of the thesis. By examining 'care left undone' we may begin to see a little more of what is happening within the 'black box' described by Mark that has nurse staffing as an input at one end, and patient mortality as an outcome at the other (Mark 1995). It focusses attention on the more proximal consequences of nurse staffing levels in terms of what care is provided or missed, and whether missed nursing care impacts on patient mortality. If care left undone not only contributes to increased risk of mortality but also mediates the relationship between RN staffing and mortality, it may point to a plausible causal pathway between RN staffing and mortality.

5 METHODS

5.1 THE RN4CAST STUDY

All four studies in this thesis draw entirely or in part on data from the 'RN4Cast' study; a large multi-country study funded through EU 7th framework (FP7 2007-2013, grant agreement no. 223468), led by Walter Sermeus and Linda Aiken, and undertaken from 2009-2011. The RN4Cast study aimed to consider the relationship between nurse staffing and a variety of nurse and patient outcomes in order to inform workforce planning models and help forecast (hence the acronym RN*4Cast*) the numbers of RNs required to deliver care to a certain standard and with certain outcomes.

A study protocol was established and followed with some adaptation by the consortium of 15 countries, 12 from the EU and three outside, taking part. My involvement in RN4Cast as the researcher with responsibility for the nurse survey in England began in 2008, meeting with other members of the RN4Cast consortium to plan and agree the study design and methodology. Working with colleagues at King's College London and University of Southampton, I led and managed the survey of nurses in England, and oversaw all aspects of data collection, data processing, preliminary analysis and feedback of the findings to the participating hospitals. Tishelman's research group at Karolinska Institutet, with which I am now also affiliated, led RN4Cast's Swedish component.

The agreed methodology and design was published by the consortium in 2011 (Sermeus et al. 2011). It builds on the approach used in the previous International Hospital Outcomes Study described in Chapter 3, funded by the National Institute of Health (Aiken et al. 2001).

To capture characteristics of the nursing workforce and the context in which hospital care is provided, including nurse staffing levels, a survey of RNs was undertaken in each participating country. The patient outcome targeted for the study was patient mortality in acute general hospitals. Patient data (such as age, gender, diagnosis, procedures) was drawn from routinely collected data. To allow the analyses to control for differences related to organisational characteristics of the hospitals, such as their size, teaching status, and technology status (based on whether open-heart and/or transplantation surgery undertaken), additional information was collected about each hospital.

5.2 ETHICS AND RESEARCH GOVERNANCE/APPROVALS

The study protocol was approved by the ethics committee at Katholieke Universiteit Leuven, which was the coordinating centre for the study (Ref: B3222009 6682), and by the relevant ethical committees in all participating countries. Country level approvals to acquire and analyse patient outcomes data were also obtained.

In England, ethical approval was sought and gained from the National Research Ethics (Ref: 09/H0808/69) and permissions acquired for the research to be undertaken at each hospital. In Sweden, ethical approval was received for the RN4Cast project from the regional Ethical Review Board in Stockholm (Dnr 2009/1587-31/5).

5.3 COUNTRY CONTEXT

Whilst the RN4Cast study involved 15 countries in total, the research in this thesis includes data from England (Study I & II), Sweden (Study III), and a dataset based on material from nine of the RN4Cast participant countries (Study IV). Study I also examined routinely collected data on staffing and mortality rates for all acute hospitals in England.

Before describing the methods, a brief description of the country contexts is provided.

In England, a National Health Service (NHS) provides care that is publically funded and free at the point of access. Following the 'provider-purchaser divide' in the 1990s, hospitals are managed and governed by 'self-governing Trusts', now simply referred to as 'NHS Trusts'. NHS Trusts are the governing bodies for what may have previously been multiple service providers: hospitals, community services, mental health services. Not all NHS Trusts provide acute services. Those that do may have more than one acute hospital, and some may also provide other services, such as community or mental health care. Nursing education is based in universities, and requires three years of clinical practice and academic study (divided equally between the two) to achieve a professional qualification as an RN (regulated by the Nursing Midwifery Council) and a bachelor's degree.

In Sweden health care is predominantly publicly funded, and is organised at three independent governmental levels: nationally, regionally (county councils/health care regions) and locally in municipalities. At the time RN4Cast was conducted there were about 79 hospitals providing acute care spread across 21 health care regions. Since 2007 nurse staffing in acute care hospitals is generally comprised of RNs (who have received a minimum of three years of education based in a university and are prepared to bachelor's level), and support nurses, who have received 18 months education in nursing. The proportion of RNs in the health care workforce increased from 57% in 1993 to 71% in 2007 (Swedish Association of Local Authorities and Regions (Sveriges Kommuner och Landsting) 2011).

The data set used in Study IV was the same set of countries that had previously been used to examine relationships between RN staffing levels and nursing educational levels and inpatient mortality following common surgery (Aiken et al. 2014). Data from nine countries were included: Belgium, England, Finland, Ireland, the Netherlands, Norway, Spain, Sweden, and Switzerland. Figure 2 shows the number of nurses (defined as 'professionally active' nurses, excluding midwives and nursing aides) per 1,000 inhabitants in each of these countries relative to other OECD countries.

Figure 2: Nurses per 1,000 inhabitants on study IV countries

Nurses per 1 000 inhabitants 2010 (OECD)
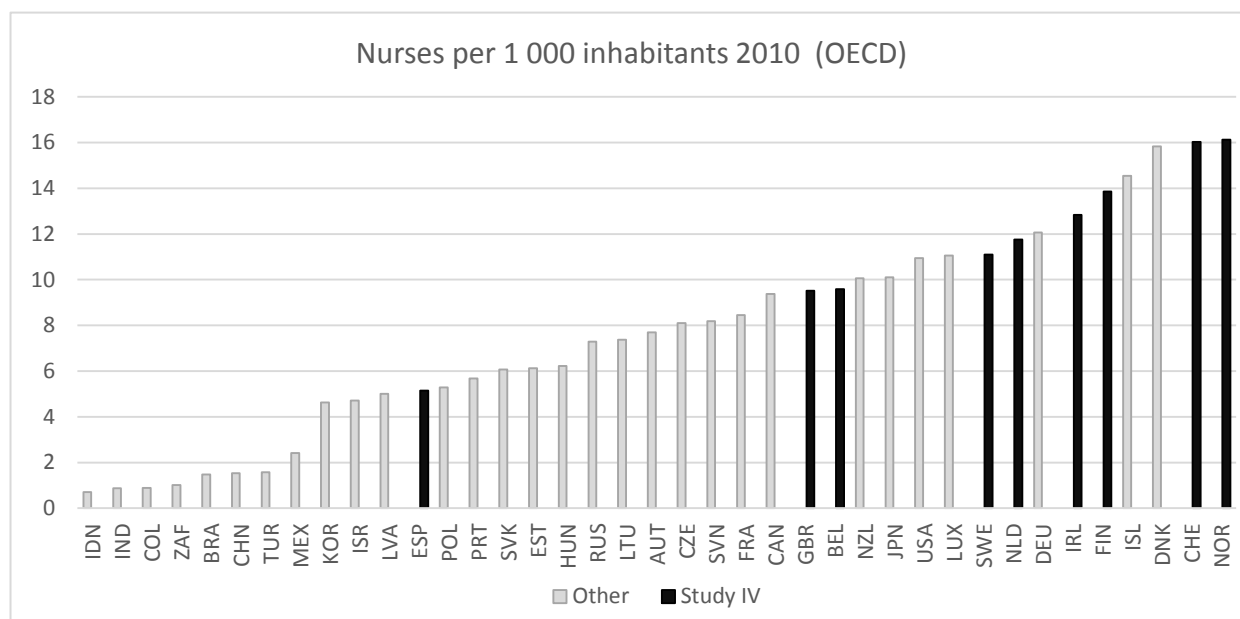
Source: OECD Health Resources - Nurses (indicator). doi: 10.1787/283e64de-en (Accessed on 05 March 2017)

5.4   SURVEY OF NURSES

The goal of the nurse survey was to obtain data from RNs on medical and surgical wards within general acute hospitals. The survey questionnaire, and measures within it, have been used in previous studies of nurse staffing and patient outcomes (Aiken, Clarke, and Sloane 2002; Aiken et al. 2002; Lake 2002)

The RN4Cast questionnaire consisted of five core sections, which were used in the surveys in all countries: 'Work Environment and Job Satisfaction', 'Quality and Safety', 'Your most recent shift', 'About you', and 'Where you work'. The key questions and derived measures are described in section 5.5.

5.4.1   Data collection: Nurse survey in England

Hospitals: The entire population of NHS Trusts with acute hospitals in England was identified. A sampling frame was set up to categorise Trusts by geographical area, size and teaching hospital status. A random stratified sample of 64 NHS general acute hospital Trusts in England was drawn to ensure mix by size, teaching status and region. All 64 were contacted by the study team, writing to the Director of Nursing and Chief Executive of each. Thirty-one of the 64 Trusts selected in the sample agreed to take part, covering 46 hospitals and 401 wards. During the course of recruitment to the study, we aimed to maintain the quotas identified in the sampling frame through proactive follow up to ensure a representative mix of Trusts was obtained in terms of size, geographic location and teaching status.

Study coordinators: Each Trust taking part nominated a designated study leader to coordinate the distribution of the survey and liaise with the research team. Typically this was an assistant director of nursing, or divisional lead nurse. In some Trusts the director of nursing undertook this role.

Wards: A maximum of 10 medical and surgical wards (five of each) in each hospital were included in the survey. All the RNs on these wards were sent a letter of invitation to the study and questionnaire (via the ward). Where a hospital had more than five eligible wards in either group we selected five at random. In Trusts that covered multiple hospitals, 10 wards from each hospital site were selected: for example, up to 20 wards from a trust with two hospitals. Based on inclusion criteria given by the research team, hospitals identified wards they considered to be 'general' medical and surgical wards. Highly specialised units and intensive care units were excluded whilst general surgical, general medical, orthopaedic, and acute care of older people were included. The final samples were checked by the study lead investigator.

Survey administration: The internal distribution of the surveys within each hospital was managed by the study coordinator, who was a member of NHS staff employed by the Trust. The data collection was staggered over eight months, from January to September 2010.

A questionnaire, free-post response envelope, RN4Cast pen, and project information sheet were issued in sealed packs, addressed to 'staff nurse' (un-named) and the name of the ward. Each questionnaire included a code that identified the Trust, the hospital and the ward, but not the individual nurse. The ward manager was also sent a letter via the internal mail, giving them further information about the study, to supplement the information they had received from the site coordinator. An online version of the questionnaire was offered as an alternative means of responding. A series of reminders were sent to the ward managers and study coordinators, including a full reminder, a postcard, and reminder emails.

Nurses: In order to send the right number of questionnaires to each hospital, and to each ward, study coordinators were asked to provide details of the wards and the number of RNs on each. These numbers also provided the denominator to allow an estimate of response rates to be calculated.

Response: Study coordinators at each Trust were sent regular updates on the response rates (breaking down the response rate by ward). Trusts with lower response rates were contacted by telephone, email and letter, to report response rates and offer support in improving response – particularly for wards where no responses had been received.

In total, 7609 questionnaires were distributed to RNs in the 46 hospital sites in the 31 Trusts. Complete responses were received from 2990 RNs, representing an overall estimated response rate of 39%. Response rates varied considerably between hospital Trusts, from an estimated 19% to 69%.

5.4.2   Data collection: Nurse survey in Sweden

The approach to achieving the core data set was different in Sweden due to the size and spread of the country. Nursing members of the nursing Swedish Association of Health Professionals were surveyed and then respondents meeting the study criteria were extracted to generate an RN4Cast dataset.

Sample: The membership of the Swedish Association of Health Professionals was used as the source to identify RNs for the survey. The register included approximately 81% of all clinically active RNs in Sweden at the time of the survey (Lindqvist et al. 2015). For each registrant, it includes information on workplace, including both department and hospital. However, it does not hold data on the RN's specific role, and whether they are working in inpatient or outpatient settings.

The survey strategy involved an intentional over-recruitment: all RNs working in medical or surgical departments were selected (N=33,083) as the population to be recruited to the survey. The use of unique national registration numbers allowed records to be shared between the union's database and a national register of residential addresses.

Translation: The questionnaire was translated into Swedish, and achieved a high content validity index of 0.91 (Squires et al. 2013).

Survey administration: The survey itself was administered by the government agency Statistics Sweden. The postal survey was sent directly to the RNs' home address during February 2010. Respondents were given the choice of completing the paper form and returning it in the prepaid envelope, or taking part through a web-based version. Three reminders were sent: at two weeks, four weeks and after six weeks. Additional printed questionnaires were included in the last two reminders.

Response & exclusions: The overall return rate at the end of the data collection period was 70% (n= 23,087). An opening question established if the respondent was part of the target population, i.e. was working actively in direct inpatient medical or surgical acute care. Two further control questions checked the hospital and department where respondents were working. Excluding cases who were not working general medical and surgical wards, a total of 10,174 respondents were included as matching the inclusion criteria used for the RN4Cast study (Lindqvist et al. 2015). These formed the dataset used for analysis in this thesis.

Note that the design of the Swedish study, did not allow the individual wards/units to be identified.

5.5   NURSING MEASURES

The questionnaire, which was the same for each participating country, comprised a number of elements that are central to this thesis. Each key component is outlined below, with details of the instruments used to collect data and the measures derived for use in the study.

Practice Environment

The 'Nursing Work Index' was modified by Aiken and colleagues for use in the USA, and subsequently used in the International Hospital Outcomes Study in USA and four European countries (Rafferty et al. 2007; Aiken and Patrician 2000). The revised instrument, termed the 'Practice Environment Scale' is used to characterise different dimensions of the nursing work environment, and comprises constituent scales which have been internationally validated in a number of studies (Lake 2002; Li et al. 2007). The index includes 32 items relating to nurse participation in hospital affairs, managerial support for nursing, resource adequacy, promotion of care quality, and views of the relationships between nurses and doctors. Respondents were asked to rate each item on a four point Likert scale (strongly disagree, disagree, agree, strongly agree). The mean score (from 1-4, where one is low) across all items was used to provide an overall rating of the practice environment.

Nurse staffing

Nurses were asked to report the total number of nursing staff giving direct patient care (specifically 'Registered Nurses' and 'Other nursing care staff') and the total number of patients on the ward during the last shift they worked. From these questions, three measures were derived:

- RN Staffing: Patients per RN providing direct care
- Nursing support staffing: Patients per non-registered nursing staff
- Skill-mix: The percentage of the total nursing team providing direct care that were RNs.

The measure of staffing used in the examination of national data from England (Study I) was based on secondary data on the number of hospital beds relative to the number of posts (full-time equivalents). The resultant measure was 'beds per member of staff'.

Role in care provision

The survey asked nurses to describe their role on the last shift they worked. They were asked to indicate if they: 'provided most care themselves', 'supervised care provided by others and provide some themselves', or 'provide only limited care with most direct care being carried out by others'.

Transferrable activity

Variation in the range of activities undertaken was measured by asking respondents how frequently ('never', 'sometimes', or 'often') they undertook nine potentially transferable

activities, such as obtaining supplies, delivering food trays, covering for other services out of hours. The list included the activity: 'performing non-nursing care' (or as translated in Swedish: 'Utförde arbetsuppgifter som borde göras av andra yrkesgrupper', literally translated: Performed work that should be done by other staff groups). The categories were combined to create a binary variable; 'often' vs. 'sometimes or never'.

Nurse education

Nurses were asked to indicate whether they held a bachelor degree in nursing. A measure of nursing educational level was produced by calculating the percentage of RNs at each hospital that reported holding a baccalaureate degree or higher in nursing.

Patient mix at the time of the last shift

Nurses were also asked to report the number of patients requiring assistance with daily living, and the number requiring frequent monitoring, to measure nursing workload intensity related to patient need.

Care quality and Patient safety

A single question asked nurses to rate the quality of care on their ward overall on a four point scale (poor, fair, good, excellent). Using an item from the Agency for Healthcare Research and Quality's hospital survey on patient safety culture, previously validated in the UK (Waterson et al. 2010), nurses were also asked to give their ward an overall grade on patient safety, with response options: 'poor', 'failing', 'acceptable', 'very good' or 'excellent'.

Care left Undone

The activities in the measure of 'care left undone' were based on a measure used by Aiken and colleagues (Aiken et al. 2001), which was further developed by Schubert et al. to form the 'Basel Extent of Rationing of Nursing Care' (BERNCA) – an instrument to assess the implicit rationing of care (Schubert et al. 2008). Nurses were asked 'On your most recent shift, which of the following activities were necessary but left undone because you lacked the time to complete them?' Respondents were presented with a list of 13 nursing care activities and asked to tick all that applied. The activities listed were:

1. Adequate patient surveillance
2. Adequate documentation of nursing care
3. Administering medication on time
4. Comfort/talk with patients
5. Develop or update nursing care plans/care pathways
6. Educating patients and/or family
7. Frequent changing of patient's position
8. Oral hygiene
9. Pain management
10. Planning care
11. Preparing patients and families for discharge
12. Skin care
13. Undertaking treatments/procedures

From these items, two main measures of 'missed care' were produced. Firstly, the reported prevalence of any care being left undone, based on one or more of the activities having been ticked (i.e. a binary measure). Secondly, a score indicating the number items of care left undone, by summing the number of activities ticked per person.

These measures were used in Study II and Study III. In Study IV, where nursing characteristics were included in the analysis at the hospital level, the total missed care score was expressed as a mean percentage across respondents, where no missed care is 0, and 13 out of 13 items missed is 100%.

5.6   OUTCOME MEASURE: MORTALITY

Mortality was the main patient outcome variable used in the RN4Cast study, and was used in Studies I and IV of my research. The multi-country study consortium agreed on a set of definitions and goals, regarding the classification of patient diagnostic and outcome data, to ensure consistency between countries. The core data set, required to enable risk adjustment, was based on data from: acute inpatients, sex, age (only aged 18+ included), ICD-10 and DRG codes (to identify primary and secondary diagnosis, procedure codes, and diagnosis related groups), length of stay data (to enable mortality within 30 days to be derived), discharge destination, and route of admission (elective, emergency).

In Study I (England), the national Hospital Episode Statistics dataset provided the source for data on all patients admitted between 1 April 2009 and 31 March 2011. The predicted number of deaths in hospital was calculated, using a technique that mirrors the approach used to produce the 'Summary Hopsital Mortality Indicator' in England (Campbell et al. 2012). This approach uses indirect standardisation for elective status, sex, age, socio-economic deprivation, a modified Charlson Index for comorbidities, and the number of emergency admissions in the preceding year (Charlson et al. 1987). Admissions were divided into medical and surgical specialties according to the specialty code of the admitting consultant. The predicted number of deaths in each group for each Trust given the case-mix was calculated.

Similarly, Study IV used inpatient mortality as the primary outcome measure; however, in this study, we focussed specifically on patients who had undergone common surgical procedures. Again, mortality was measured using administrative data on discharge status (death or survival), length of stay (less than 30 days and adjusted for surgical procedure), patient age, sex and admission type (emergency or elective), and comorbidities (Silber et al. 2000). The Charlson comorbidity index was once again used to control for differences between patients on admission (Charlson et al. 1987).

5.7 ANALYSIS

For each of the studies, statistical analyses (using a variety of techniques and statistical software packages) were undertaken in collaboration with, and under the supervision of, expert statisticians. The regression models were used to examine: the associations between nurse staffing and case-mix adjusted mortality whilst taking account of other factors (Study I), what part care left undone plays in this relationship (Study II and IV), and how the context in which nursing care is provided influences the relationships between staffing and care left undone (Study III).

Multivariate regression models were used to explore the relative chances of a particular outcome, related to different levels of the factors considered. Multilevel models were used to account for the nested hierarchical structure of the data: for example, nurses within wards/units, within hospitals, and within countries.

The level that each variable pertains to is stipulated in the analytical model. So, for example, the specialty (medical or surgical) and practice environment were identified as ward level variables in the model exploring care left undone and staffing in England (Study II), whilst the time of shift or mix of patients were identified at the individual shift level. Such models also allowed the amount of variance occurring at each level to be described.

5.8 METHODS FOR EACH STUDY

Table 1 summarises the key variables, data sources and samples for each of the four studies.

Table 1. Summary of data sources and samples for the four studies

Study	Topic/key variables	Country	Data Source	Unit of analysis	No. Hospitals	No. Nurses	No. Patients
Study I	Staffing (RN, support worker, medical) & Mortality	England	a) National	All admissions	137 (all acute Trusts)	-	9,669,555 Medical admissions 9,302,292 Surgical admissions
	Staffing (RN, support worker, medical) & Mortality	England	b) RN4Cast	All admissions	46 hospitals (31 Trusts)	2,990	
Study II	Staffing (RN, support worker) & Care Left	England	RN4Cast	Nursing shifts	46 hospitals (31 Trusts)	2,917	-
Study III	Other Factors, Staffing (RN, support worker) & Care Left	Sweden	RN4Cast	Nursing shifts	79	10,174	-
Study IV	RN Staffing, Mortality & Care Left	9 EU countries	RN4Cast	Patients following common surgery	300	26,516	422,730

6 SUMMARY OF THE STUDIES

The full manuscripts arising from each of the four studies are presented in the second half of this thesis. In this chapter the studies are summarised and a narrative is provided to make explicit how the studies connect to one another, the key findings from each, and how the thesis evolved from one study to the next. The chapter ends with an overview of key findings across the studies.

6.1 STUDY I

Griffiths P, Ball J, Murrells T, Jones S, Rafferty AM. Registered nurse, healthcare support worker, medical staffing levels and mortality in English hospital trusts: a cross-sectional study. BMJ Open. 2016 Jan 1;6(2):e008751.

As described in Chapter 3, the first of the hypotheses – that RN staffing relates to mortality rates in hospitals – has been frequently tested and the association has been confirmed in research that spans decades. However, there are few examples of studies that simultaneously examine the contribution of other staff, such as medical doctors and health care support workers. Study I thus sought to address this gap in the literature, and examine the relationship between nurse staffing levels taking account of the presence of other staff.

More precisely, the aim of the first study was to examine the association between RN staffing levels and patient mortality rates in acute hospitals, whilst also examining the staffing levels of doctors and health care support workers (also referred to as nursing support staff).

Two separate sets of data were drawn upon. The first was an analysis of secondary data from all 137 NHS general acute hospitals in England, exploring the relationship between workforce and hospital standardised mortality rates for adult medical and surgical admissions in 2010 and 2011. Regression models were used to explore the main effects and interactions between the staff groups. The second set of data applied similar regression models to a subset of acute hospital trusts: the 31 RN4Cast NHS Trusts in England survey. This allowed a more detailed workforce measure to be examined in relation to mortality rates.

Differences in both medical and nurse staffing levels were found to be associated with differences in hospital mortality rates, in both the full national dataset and the RN4Cast subset of hospitals. However, when both nurse staffing and medical staffing were included within the same case-mix adjusted regression models, the effect of medical staffing was dominant. The effect of RN staffing was no longer significant in the adjusted models using the national data for surgical or medical admissions. It remained significant in the adjusted model for medical admissions in the subset of RN4Cast hospitals; relative risk of mortality was 11% less for

patients in hospitals with average ward staffing of six patients or less per RN, compared with those with 10 patients or more (RR0.89, p=0.001). An association in the opposite direction was found in relation to support worker staffing, for medical admissions: higher support worker levels were associated with lower mortality rates (RR 0.93, p=0.003). Overall, compared with previous studies, this study revealed a less clear picture of associations between RN staffing and mortality.

However, the study had limitations that need to be considered in interpreting the findings. Neither of the datasets used (the national secondary data for all acute hospitals in England, and the RN4Cast study data from a sample of 31 acute hospital Trusts) included *both* detailed measures of nurse staffing *and* measures of medical staffing. Each dataset represented a trade-off between depth and breadth. The data in the national analyses, whilst based on all acute hospitals in England (so having a larger sample size and greater statistical power), used a relatively crude measure of staffing for both staff groups: total posts per bed. The RN4Cast data, using a subset of hospitals in England, included a more sophisticated ward-level nurse staffing measure: direct reports of nurse to patient ratios per shift. However, the same relatively crude measure of 'bed to post' ratio was used to assess medical staffing levels. Arguably the biggest limitation however, was the small sample size relied upon in this subset sample: 31 Trusts (46 hospitals) all drawn from a single country. Patient deaths in acute hospitals are rare, and deaths that can be attributed to characteristics of hospital care – referred to in some research as 'avoidable mortality' – are rarer still. The lack of sensitivity of this outcome measure had serious implications for the sample size needed to power the analysis.

Despite its limitations, the study nonetheless offered some, albeit weak, support for the previously observed patterns of association between RN staffing and mortality rates. It did not, however, offer any further insight into possible mechanisms for the relationships reported between different measures of staffing and patient mortality in acute hospitals. Plausible causal pathways remained untested.

6.2 STUDY II

Ball J E, Murrells T, Rafferty AM, Morrow E, Griffiths P. 'Care left undone' during nursing shifts: associations with workload and perceived quality of care. BMJ Quality & Safety. 2014;23(2):116-25.

This study examined the more proximal potential consequences of RN staffing levels on the care that RNs do (or, more precisely, do not) deliver, and the implications for the quality of that care. Specifically, the study sought to use data from the nurse survey in England in order to determine if there is a relationship between nurse staffing levels and necessary care left undone. It was hypothesised that lower nurse staffing levels may result in RNs having less time to complete necessary nursing care, and thus in higher levels of care left undone per shift. Within

the RN4Cast nurses survey, RNs were asked to report, through a closed question offering 13 responses, which, if any, necessary nursing care activities had not been done on their last shift due to lack of time. The survey also captured nurses' assessments of the quality of care where they worked and asked nurses to rate the patient safety of the environment.

The study used data from the RN4Cast nurse survey in England; a cross-sectional survey of 2,917 RNs working in 401 general medical/surgical wards in 46 general acute NHS hospitals in England.

Simple frequencies revealed that 86% of RN respondents had left one or more care activities undone due to lack of time on their last shift. The activities most frequently left undone were: comforting or talking with patients (66%), educating patients (52%), and developing/updating nursing care plans (47%). Using multilevel regression models, we found that the RN staffing levels at the time of the shift were significantly associated with the incidence of 'missed care' ($p<.001$), even when controlling for HCSW staffing levels and differences between wards and hospitals. The size of the effect was most visible in comparing the best-staffed quintile (six patients or fewer per RN) with the others; the top 20% of ward shifts in England had staffing sufficient to significantly reduce the amount of care being left undone by RNs (RR -1.087, $p<0.001$). HCSW levels did not reduce the likelihood of care being left undone by RNs.

We also found a relationship between care left undone and nurses' assessment of safety and quality. On average, 7.8 activities per shift were left undone on wards which were rated as 'failing' on patient safety, compared with 2.4 on those where patient safety was rated as 'excellent' ($p<001$). These findings confirmed an association between care left undone and nurse rated safety scores. The same sort of relationship was found in relation to quality. A potential limitation, however, is that these measures are self-reported by RNs.

Study II established that when nurse staffing levels are low, more necessary nursing care is reported to be left undone due to the lack of time. This is far from being a scarce problem on English NHS hospital wards and our study found a statistically significant relationship between RN staffing and care left undone.

The findings pointed to a plausible theory that nurse staffing levels could be related to patient outcomes, because of the increased risk of necessary nursing care being left undone when RN levels are lower. However, this paper does not tell us the extent to which the relationships reported are specific to the England context.

6.3 STUDY III

In Study II a relationship was found between RN staffing levels and care left undone on hospital wards in England. In Study III we sought to establish if the same relationships can be found in Sweden, in which the context of care potentially differs. The following aspects of care context were explored in relation to care left undone and staffing: the role of RNs in supervising or delivering care, the staffing deployment patterns across 24 hours, and the tasks delegated to others or retained by RNs.

The relationship between 'care left undone' and 'staffing' was again examined using multivariate regression techniques, this time including a wider range of independent variables and potential confounders. However, the key to this study was to not simply replicate the analysis of Study II – and focus on the relationship between 'care left undone' and staffing, whilst controlling for the impact of other variables – but to also explore the impact of 'other' factors on staffing, care left undone and the relationship between the two. The aim of the paper was therefore to consider the nursing care delivery 'context'; to explore the way in which some of these 'other' factors have a relationship with one another and with the variables of interest, and to provide a descriptive account of these relationships.

Specifically, the study examined 'care left undone' by RNs on medical/surgical wards in acute care hospitals in Sweden and explored factors associated with variation in amounts and types of missed care. The study used data from the RN4Cast nurse survey in Sweden, and is thus based on a cross-sectional survey of 10,174 RNs working on inpatient general medical/surgical wards in 79 acute care hospitals (Jan-March 2010). A multi-level (nurses nested in hospitals) logistic regression model was used to examine the effect of key factors, including nurse staffing, on 'care left undone'.

The first step in the analysis was to establish and present a descriptive picture of some of the contextual factors.

Three-quarters (74%) of RNs reported that they left some care undone on their last shift. Staffing levels varied by time of shift with early shifts being better staffed than afternoons (mean 5.5 patients per RN vs 7.0 patients per RN) and both being much better staffed than nights (11.4 patients per RN). There is some difference in the activities undertaken according to the time of shift, with nurses on afternoon shifts being less likely to have arranged patient discharges but more likely to have answered the phones or filled in for non-nursing staff. Taking account of the many factors relating to care left undone (shift length, role, patient mix,

practice environment), the odds of care being left undone is halved on shifts with RN staffing levels of six or fewer patients per RN compared with on shifts of 10 or more patients per RN (OR 0.466, p<0.001). Support workers were found to have a beneficial effect only when comparing the furthermost extremes: the odds of care being left undone on units that had no support workers compared to those that had the most support workers (OR 0.714, p=0.021).

In Sweden, as in England, we found an overall association between RN staffing and care left undone. Exploring the care delivery context offered some new insights into the circumstances in which care that is considered necessary is nonetheless left undone. Whilst necessary care being left undone may be of interest in its own right, neither Study II nor Study III helped to establish if care left undone is correlated with an objectively reported measure of patient outcome, such as mortality.

6.4   STUDY IV

> Ball J E, Bruyneel L, Aiken L, Sermeus W, Sloane D, Rafferty AM, Lindqvist R, Tishelman C, Griffiths P. Post-operative mortality, missed care and nurse staffing in nine countries. *Submitted for publication*.

Studies II and III established that there is a relationship between RN staffing levels and care left undone, and that this relationship exists in multiple contexts regardless of specific deployment differences. This finding was corroborated by further analyses based on data from the RN4Cast study for 488 hospitals in 12 countries, to which I contributed (Ausserhofer et al. 2014).

Although limited by sample size, Study I used RN4Cast data to examine the relationships between case-mix adjusted mortality and staffing level variables. In 2014, analysis of nurse staffing variables from the RN4cast study for nine countries – as opposed to the single country analysis used in Study I – reported a statistically significant relationship between staffing and mortality rates for patients (Aiken et al. 2014).

Study IV focussed on the final part of the theory underpinning this thesis, to see whether care left undone mediates (partially or wholly) the relationship between RN staffing and patient mortality in acute hospitals.

The study built on the approach previously used by the RN4Cast consortium to examine the relationship between nurse staffing factors and hospital related mortality for patients following common surgical procedures, using the nine-country dataset (Aiken et al. 2014). The aim of Study IV was to examine associations between 30-day inpatient mortality, RN staffing and nurse reports of 'care left undone'. The dataset was comprised of routinely collected data on 422,730 surgical patients, from 300 general acute hospitals, and a survey of 26,516 RNs.

A generalized estimation approach with random intercept models was used to examine the relationship between first staffing, and then 'care left undone', and mortality. Bayesian methods were used to test for a mediating effect. The analysis followed four steps in order to test the individual associations and the mediating role of care left undone, focusing on:

1. Staffing and mortality (controlling for hospital and patient factors)

2. Care left undone and mortality (controlling for hospital and patient factors)

3. Staffing on care left undone and on mortality (controlling for hospital and patient factors, plus other nurse factors – practice environment, nurse education)

4. Staffing and care left undone (simultaneously) on mortality

The results show that 'Care left undone' was significantly associated with 30-day case-mix adjusted inpatient mortality. A 10% increase in the percent of 'care left undone' was associated with a 16% (OR 1.159 95% CI 1.004-1.026) increase in the likelihood of a patient dying within 30 days of admission following common surgical procedures. When RN staffing and 'care left undone' are both in the same model, the effect of staffing on mortality is reduced, signaling a mediating effect. The study does not however explore the impact of different aspects of missed care on risk of mortality in hospital. It should also be noted that the study examines the outcomes of patients who had undergone common surgery in particular; we do not know if these findings would apply to other patient groups.

The results support a potential explanation for the frequently observed association between nurse staffing and mortality: where RN staffing levels are lower, nurses do not have time to complete all necessary care, and that care left undone or missed increases the risk of patient mortality in hospital, taking other factors into account.

6.5   SUMMARY OF KEY FINDINGS

The key findings from the research in this thesis can be summarised as follows:

- Studies examining nurse staffing and patient outcomes have tended to ignore medical staffing. Using data from England, we included both support worker and medical staffing levels in a model with RN staffing and case-mix adjusted mortality. Whilst the findings were more equivocal than some studies, we found an association between ward-based RN staffing level and mortality rates of medical patients (but not surgical) in models that also included medical staffing and support worker staffing. (Study I).
- We found no evidence of a positive effect on mortality rates associated with higher levels of nursing support staff – but some evidence that higher nursing support staff levels were associated with worse patient outcomes. (Study I).

- Necessary nursing care is frequently reported by regsitered nurses (RNs) as being left undone due to lack of time; 86% of RNs surveyed in England and 74% in Sweden reported that they left some care undone on their last shift. (Study II, Study III).

- Nurse-rated quality of care and nurse-rated patient safety scores were both significantly more likey to be positive when reported levels of care left undone were lower. (Study II).

- When RN staffing levels were lower, there was an increased likelihood that necessary nursing care was left undone. (Study II, Study III)

- Higher levels of less skilled nursing staff (health care support workers or nursing aids) were not associated with lower levels of care reported as being left undone by RNs in England (Study II). Support workers were found to have a marginally beneficial effect in Sweden (Study III).

- Despite some differences in the care context between England and Sweden, the same fundamental relationship between care left undone and RN staffing was found in both countries, and in an analysis of data from nine EU countries. (Study II, III, & IV).

- In a nine country analysis focussed on outcomes following common surgical procedures, care left undone was significantly associated with case-mix adjusted 30-day mortality. In this model, care left undone mediates the relationship between RN staffing and patient mortality and offers a plausible causal pathway between RN staffing levels and patient mortality. (Study IV).

7 DISCUSSION

The overarching aim of the research in this thesis has been to systematically investigate relationships between RN staffing levels, care provision – or, more precisely, nursing care that is left undone – and the risk of patients dying as a result of hospital related factors. Using an observational design with cross-sectional data, the research in this thesis examined RN staffing and patient mortality, concurrently with medical and support worker staffing. It then went on to study the part played by care left undone as a potential consequence of nurse staffing, other factors associated with care left undone, and care left undone as a predictor of patient mortality.

The research in this thesis has tested the hypotheses that lower RN staffing levels are related to a higher risk of care being left undone, and that higher levels of care left undone are associated with higher levels of case-mix adjusted hospital mortality.

Here in the discussion, key findings are reviewed in turn to explore how the research in this thesis relates to previous research, and what it adds beyond what was previously known. Whilst specific limitations of the research are outlined in relation to each key finding, a broader review of methodological issues is then presented, before considering the implications for future research. The chapter ends with a consideration of the implications of the findings for policy and practice.

7.1 DISCUSSION OF KEY FINDINGS

7.1.1 Higher RN & medical staffing levels are associated with lower risk of patient mortality

Higher medical staffing is significantly associated with reduced mortality for medical and surgical patients in English hospitals. Higher ward-based RN staffing is significantly associated with lower mortality for medical patients. The effect of RN staffing on mortality is attenuated in models that included medical staffing. (Study I)

The role boundaries staff groups are fluid: there is potentially scope for the work of one group to substitute for that of another. Yet, despite literature spanning 30 years that has examined nursing workforce and patient outcomes, very few studies have looked at the registered nursing workforce in conjunction with other staff groups, namely medical staffing and nursing support workforce. Failure to consider staff groups simultaneously has been a significant limitation in much of the research that seeks to understand the relationship between workforce and patient outcomes in hospitals.

It might be supposed that the lack of research examining the medical and nursing workforce concurrently is due to the literature being split along disciplinary lines, with research on nursing workforce sitting in a separate disciplinary research space to that on medical workforce.

However, searching the literature reveals just a few published studies that have examined medical staffing levels in relation to patient outcomes (Bond et al. 1999; Jarman et al. 1999; West et al. 2014).

Our research found that differences in both medical and RN staffing levels were associated with differences in hospital mortality rates, in both the full English national dataset and the English RN4Cast subset of hospitals. However, when both nurse staffing and medical staffing were included within the same case-mix adjusted regression models, the effect was clearer for medical staffing. The effect of RN staffing was no longer significant in the adjusted models using the national data for surgical or medical admissions. It remained significant in the adjusted model for medical admissions in the subset of RN4Cast hospitals.

The predominance of medical staffing as a predictor of standardised hospital mortality rates confirmed findings from the few previous studies that included both. Examining data for 183 acute hospitals in England between 1991-1995, Jarman and colleagues reported that the number of doctors (per hospital bed, and general practitioners per head of population) were significantly associated with standardised hospital mortality ratios (Jarman et al. 1999). The only nurse staffing variable which added to the predictive power of the model was the proportion of the lowest grade of nursing support staff in the workforce (higher levels of support staff were associated with higher levels of standardised hospital mortality). In a study using data from 3763 United States hospitals from the same era, which examined the case-mix adjusted mortality rates in relation to staffing levels (of a range of health care staff), a significant relationship was found between RN staffing and mortality, even after adjusting for medical staffing (Bond et al. 1999).

More recently, West and colleagues examined both nursing and medical staffing in ICUs in England. In this study, where the setting and client group were limited to a single type of unit rather than across whole hospitals, researchers found that both medical staffing and RN staffing levels had a significant effect on case-mix adjusted patient mortality (West et al. 2014).

A strength of the national analyses in Study I is that being based on all 137 acute hospital Trusts in England, there is a larger sample size and greater statistical power than the subset of 31 RN4Cast hospital Trusts. However, a limitation that applies to both the national and RN4Cast datasets, is the reliance on administrative data for medical staffing, and the resultant lack of granularity. The measure used provides a proxy for achieved staffing levels by looking at the number of beds per post (that is the total number of hospital beds divided by total 'whole time equivalent' of staff employed). This dataset, which is the only source of publicly available staffing data for NHS Trusts in England, does not identify the portion of staffing that is related to the provision of general acute inpatient care and related to the denominator: hospital beds. The size of this compromise, and the extent to which it applies equally between staff groups and hospital trusts, is not known. It is unclear, therefore, the extent to which the 'beds per post' measure offers a reliable proxy for staffing levels on acute wards within the hospital.

We get some idea of the scale of the staffing measurement issue when we compare the variation in nurse staffing based on 'beds per post' to the nurse reported measure of 'patients per RN' used in the RN4Cast study. The beds per post measure varied from a hospital mean of 0.69 beds per RN, to 2.81 beds per RN: a four-fold difference between the minimum and maximum. In contrast, a roughly two-fold difference was found between the minimum and maximum RN staffing as measured at ward level: e.g. a minimum of 4.9 patients per RN to a maximum of 11.1 patients per RN on medical wards.

An added complication in assessing medical staffing levels, is that there is substantial variation in medical skill-mix within England. A national review by the Audit Commission found a six-fold variation in the number of junior doctors per consultant and a more than two-fold variation in number of admissions per doctor (Audit-Commssion 2010). Our measure of nurse staffing takes some account of nursing skill-mix by measuring RN staffing levels separately from that of the nursing support workforce, yet the skill-mix of medical staffing is not disaggregated.

The staffing data from the smaller subset of hospitals in the RN4Cast study provided a more precise ward-level measure of nurse staffing, using direct reports of the number of RNs and nursing support staff on duty and the number of patients present, to compute nurse to patient ratios per shift. However, the analysis of the RN4Cast subset still relied upon the 'beds per post' measure to assess medical staffing levels.

The analytical steps taken to case-mix adjust the mortality outcome are described in detail in Study I. But in common with other studies that use patient level risk factors to adjust for case-mix within the analytical model, a single resultant variable of 'case-mix adjusted mortality' is not generated, and therefore not described. Variation in the level of 'case-mix adjusted mortality', that is the portion that may be related to non-patient factors, and be reflective of hospital differences in care, cannot therefore be compared between studies, between hospitals, or between countries.

Despite these methodological challenges, we nonetheless detected some association between staffing levels and patient mortality, as described above. The study cannot elucidate whether the predominance of the medical staffing effect is because it is the more critical factor in determining patient survival or death in hospital, or whether it is an artefact of the measures used. The question of interest to this thesis was not however, which variables explain most variation in mortality, but whether some variation in mortality is related to RN staffing, even when other staffing contributions are controlled for in the statistical model. This study has found evidence that the relationship between nurse staffing and mortality remains even after control for medical staffing levels.

7.1.2 Necessary nursing care is left undone

It is not uncommon for necessary nursing care to be left undone by RNs on a shift due to lack of time; 86% of RNs surveyed in England and 74% in Sweden reported that they left some care undone on their last shift.

Study I sought to address a weakness in the evidence base – potential confounding related to the input of other staff – by examining the relationship between nurse staffing and mortality while controlling for medical staffing. The finding that there is a relationship between staffing and mortality, even when controlling for medical staffing, contributes to a sizeable body of evidence that reports on the association between RN staffing and patient outcomes (Kane et al. 2007; Griffiths et al. 2014; Shekelle 2013b). Relatively little research attention however has examined why such a relationship might exist and the mechanisms through which it may operate, although care left undone (also termed missed care) is suggested as a potential mediator between RN staffing and patient outcomes (Needleman 2016).

The findings from Study II and Study III described the prevalence and frequency of nursing care activities being left undone due to lack of time. Aside from the overall prevalence of missed care, Study II and III reported on the relative prevalence of each of 13 activities being missed. The activities most frequently left undone in England were: comforting or talking with patients (66%), educating patients (52%), and developing/updating nursing care plans (47%).

In Sweden, the care activities most frequently left undone were: comforting or talking with patients (46%), developing/updating nursing care plans (34%) and oral hygiene (31%). Pain management and treatment procedures were least likely to be left undone in both countries (6% for each in Sweden, and 7% and 11% respectively in England). (Study II, Study III).

The relative frequency that each of thirteen different care activities were reported as being left undone broadly followed the pattern described by Jones, in a review of studies of missed care; communication and emotional support activities were more likely to be left undone, whilst clinical activity was less frequently neglected (Jones, Hamilton, and Murry 2015).

Theoretical frameworks put forward to explain the association between RN staffing and patient outcomes have placed patient surveillance as a key mechanism: when staffing levels are lower, there is less adequate surveillance of patients, which results in fatal deterioration going undetected. Thirty-five percent of RNs in the survey in England and 20% of RNs in Sweden reported that 'adequate patient surveillance' was an aspect of care on their last shift that they had missed.

A limitation of the measures of care left undone in this and other studies of missed care, is that it is captured through a self-reported measure. The formulation of the question used concatenates several questions into one: if care is missed, if it was necessary, and if the reason for missing it was related to lack of time. By using a simple tick box response, the instrument used creates an abstract measure: number of activities missed on the last shift. The 'size' of

each of these units of missed care, in terms of time each activity would have taken, or the significance to patient care, or frequency that it has been missed, is not captured. The strength of this simple design is that it is easy to answer and can be applied in multiple settings as part of the larger RN4Cast questionnaire.

Based on this simple means of capturing care that is left undone, we have established that the majority of RNs surveyed, more in England than in Sweden, had left some necessary care undone on their last shift as they lacked the time to complete it.

7.1.3 Care left undone is related to RN staffing levels

When RN staffing levels are lower, there is an increased likelihood that necessary nursing care is left undone. (Study II, Study III)

This finding, reported from 46 acute care hospitals in England (Study II) and all 79 acute care hospitals in Sweden at the time of the study (Study III), corroborates findings from previous research studies: lower RN staffing levels are associated a higher risk that care is left undone.

In a review of the literature on missed or unfinished care, 'inadequate labour resources' is consistently identified as the most significant contributing factor to unfinished care (Jones, Hamilton, and Murry 2015). That there is a relationship between care left undone and staffing is thus not an unexpected finding. What distinguishes Studies II and III from other studies however, is that they are larger scale and have taken account of several care context characteristics through multilevel, multivariate regressions models. They also offer a greater level of specificity about the risk of care left undone in relation to different levels of RN staffing, in two specific countries with different healthcare and organisational systems.

During the day, RN staffing levels were typically better on hospital wards in Sweden than in England: 5.9 patients per RN on day-time shifts in Sweden compared with 7.8 in England. In both cases, however, the chances of care being left undone were significantly greater on shifts with the lower levels of RN staffing. The odds of care being left undone on medical/surgical wards in Sweden is three times greater on shifts with RN staffing levels of 10+ patients per RN compared with shifts with less than six patients per RN (OR 0.348, p<0.001).

Similarly, in England, the odds of care being left undone were three times greater when comparing the bottom quintile of shifts, where staffing levels were 11.67 or more patients per RN, with the top quintile, with 6.13 or fewer patients per RN, (OR 0.343, p<0.001).

Despite differences between the countries in their average RN staffing levels and the prevalence and frequency of care being left undone, the relationship between RN staffing levels and the amount of care being reported as left undone follows a similar pattern, as the bar chart in Figure 2 illustrates.

Figure 2. Country comparison of mean number of care activities left undone on a day shift, by RN staffing level
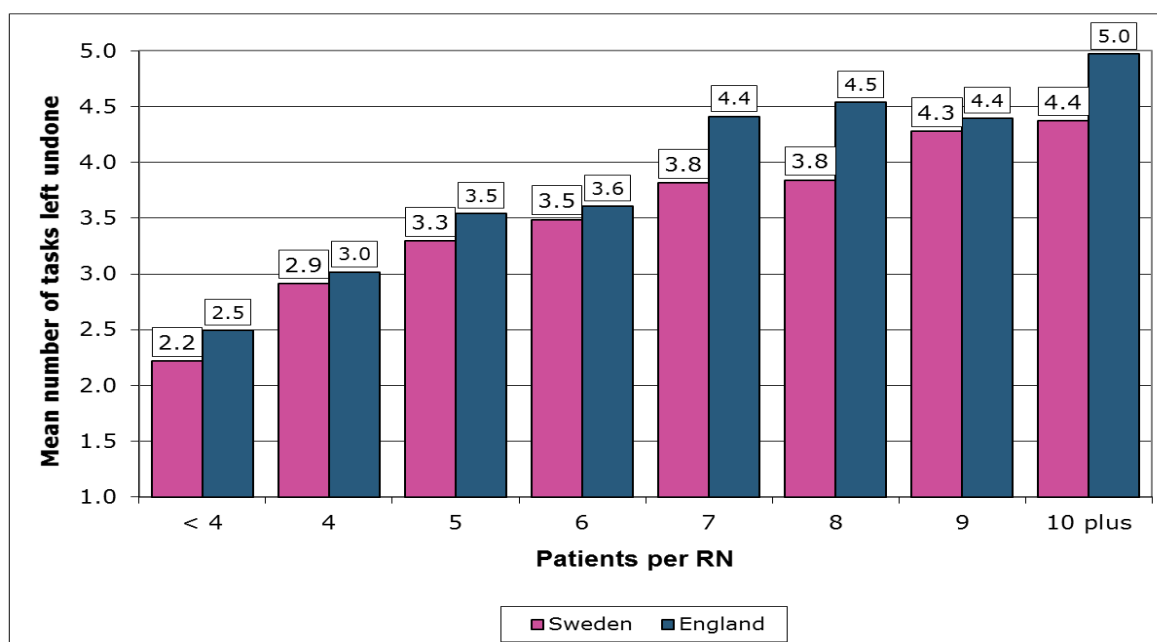
Whilst we see a significant association between RN staffing and the total care left undone measure, there are some interesting differences in the relationship to RN staffing to particular aspects of care. Although the majority of items exhibited a statistically significant association with RN staffing levels, five did not: treatments & procedures, frequent changing of patient position, administering medications on time, pain management, and preparing patients & families for discharge. This may support an interpretation that the activities missed reflects active prioritising by RNs; some care, particularly the more clinical aspects of care, are not compromised, regardless of the staffing context on the shift.

We should note that these findings are based on a measure of care left undone which presents nurses with a set list of care activities, which are applied regardless of what constitutes 'normal nursing care' activities in different settings/countries. Other aspects of nursing care, that may potentially be more relevant in one country than another due to differences in the role of the RN in different places, may not be detected. Nonetheless using a standardised measure has the advantage of having allowed data to be captured consistently in both Study II and Study III, and across RN4Cast countries, in a way that allows the relationship with staffing to be explored.

In Study II the effect of RN staffing on care left undone was strongest for 'adequate patient surveillance', which was twice as likely to be reported as left undone on shifts where RN staffing was in the bottom quintile (11.67 patient per RN) versus the best staffed shifts (less than 6.14 patient per RN). A similar pattern was found in Sweden: during a day shift RNs who reported that they had not been able to provide adequate patient surveillance were on shifts with an average of 6.7 patients per RN, compared with those that did not report this as undone, where the average was 5.7 patients.

7.1.4 Higher support worker staffing levels are not associated with better outcomes

Higher levels of nursing support staff (healthcare support workers or nursing aides) are not associated with lower levels of care being left undone by RNs in England: they neither complement nor substitute the work of RNs. (Study II, Study III).

There is evidence that in some contexts higher support staff levels are associated with an increased risk of hospital related mortality. (Study I)

One methodological challenge, discussed in more detail later, is how measures of RN staffing and nursing support staffing have been constructed and used in analyses. In the research undertaken for Studies I, II and III, the measures of RN staffing levels and support worker staffing levels were defined independently from one another: the contribution of each to the risk of patient death or of care left undone could therefore be reported separately.

The presence of greater numbers of nursing support staff was not however associated with a significant difference in the likelihood of RNs completing care or reporting that some had been left undone (Study II). The interaction between patients per RN and patients per support worker, examined in Study II for the nurse survey data from England, was also not statistically significant. In Study III, the relative risk of care being left undone by RNs in Sweden was not related to an increase in support worker levels, except at the furthest extremes in staffing: 10 patients per support worker versus less than four patients per support worker.

These findings refute the supposition that nursing support workers can 'free up RN time' (Merrifield 2016). Although substitution for nurses by assistant nursing personnel has been proposed as a means of improving the efficiency of care delivery (Spilsbury et al. 2011; Richardson et al. 1998), we found no evidence that healthcare support workers were acting as substitutes for RNs in England. Nor did we find evidence that the availability of nursing support staff increased the ability of RNs to complete their work, i.e. acted as 'complements' to RNs. Whilst we did not study all the potential work of RNs, this finding casts doubt on the premise that increasing the number of nursing support staff would increase the efficiency of RNs.

Meanwhile Study I found that increased levels of nursing support staff did not have a positive/preventative effect on the risk of patient mortality, and a detrimental effect was found for medical patients, in the analysis of the national dataset. The findings from Study II in relation to the risk of RN work being left undone – that nursing support staff neither substitute for RNs (i.e. act as replacements) nor do they act as complements to RNs (i.e. enable RNs to be more productive) – may offer an explanation for the finding from Study I. The observed relationships are congruent with the hypothesis that care left undone mediates the relationship between RN staffing and patient mortality in hospital.

This is a key finding in this thesis. There has been no shortage of studies that have set out to examine the impact of nursing skill-mix on care delivered or on patient outcomes (Griffiths et

al. 2014; Kane et al. 2007; Lankshear, Sheldon, and Maynard 2005). However, whilst the overall message from multiple studies is that a richer skill-mix is associated with better outcomes, the specific effect of the level of support workers, as opposed to the relative proportion, has been difficult to identify in the evidence base.

Nonetheless, the direction of the relationships found in the thesis research align with findings from previous studies; both those where nursing support staffing is examined in relation to RN staffing, as 'skill-mix' (Estabrooks et al. 2005; Blegen et al. 2011; Shekelle 2013a) and the rarer studies where support worker staffing is examined concurrently with RN and medical staffing (Bond et al. 1999; Jarman et al. 1999; West et al. 2014).

Since Study I was published, the issue of skill-mix and mortality rates has been examined using the nine-country RN4Cast data set, using ward-level estimates of shift staffing, and focussing on 30-day case-mix adjusted mortality for a more tightly defined group of patients: those who had undergone common surgical procedures (Aiken et al. 2016). The new analysis supports the findings from the research in this thesis: a richer nurse skill mix was associated with lower odds of mortality following common surgery, or conversely, that where greater reliance was made of support workers, risk of mortality increased.

7.1.5 Care left undone correlates with quality & patient safety scores

Nurse-rated quality of care and patient safety environment scores are significantly related to differences in care left undone. (Study II)

Arguably, reports of necessary care being left undone are of interest in their own right. It is reasonable that patients and the public expect necessary nursing care to be provided to patients in hospitals, and that they will be kept safe from harm. This expectation is eschewed, for example in the NHS Constitution in England and in the patient safety legislation introduced in Sweden in 2011 (Ridelberg, Roback, and Nilsen 2014). We therefore expect that necessary care being left undone would mean that care is of a poorer quality. This expectation is supported by the findings from Study II.

A strong relationship was found between the number of items of missed care and nurses' perception of quality of nursing care, as well as nurses overall grading of patient safety on their unit/ward. For example, the average number of care activities reported as having been left undone on wards rated as having 'poor' quality of care was 8.1, compared with an average of 2.78 where quality of care in general was rated as 'excellent'. It is worth noting that nurses' reports of care quality and patient safety are significant; they have been correlated with a number of objective measures of patient outcome such as case-mix adjusted mortality rates (Smeds-Alenius et al. 2016; Sochalski 2004; McHugh and Stimpfel 2012), suggesting a possible link between missed care and mortality rates.

The relationship between care quality measures and care left undone underscores the importance of nurses being able to deliver care completely (without leaving necessary care undone) to the overall quality of care on a ward and to the patient safety environment. However, we cannot discern from these data the direction of this association: does more missed care cause lower safety ratings, or is working in an environment with poorer safety ratings more likely to lead to nurses leaving care undone? Arguably the direction of a link between care left undone and care quality, and patient safety ratings, may not be important.

The fact that missed care, as measured through this scale, is closely correlated to judgement of the overall quality and safety is important in confirming it as an indicator of care compromise.

7.1.6 Care left undone – context

The extent to which care is left undone is shaped by factors other than RN staffing levels. The relationship between care left undone and RN staffing is found in multiple contexts, despite differences in the context of care delivery. (Study III)

One of the challenges that have been identified, in using the research evidence on nurse staffing to inform practice and policy, is that of generalisability (Griffiths et al. 2014). For example, when NICE set out to develop guidelines on safe nurse staffing levels for acute wards in England, few studies – with the exception of that from Rafferty et al. – were based on English data (Rafferty et al. 2007). And, aside from papers from the RN4Cast study (Aiken et al. 2016; Smeds-Alenius et al. 2016) we have found no research on nurse staffing levels and mortality from Sweden. Those responsible for producing guidelines and generating national policy on nurse staffing rightfully question the extent to which the findings from one healthcare system, and particular context, can be applied to another; in other words, the external validity of the findings (Bowling 2014).

The effort to understand contextual differences in order to ensure generalisability creates challenges in international research. In single country studies, the country context and, as a result, some of the contextual care delivery characteristics that differ between places, such as the type of support worker use and level of supervision and regulation, are relatively static. For example, in Kalisch et al.'s studies of missed care in Michigan (Kalisch, Tschannen, and Lee 2012), in my description of care left undone in England, or Schubert's studies on implicit rationing in Swiss hospitals (Schubert et al. 2009), country specific factors are relatively consistent across the samples drawn. A consequence of these background factors being relatively stable, is that they stop being visible within a single country analysis - for example, all hospitals in the England study are publicly funded within the NHS. The extent to which the findings from a study in one country could apply to another can be unclear.

Multi-country studies, that present analysis of RN4Cast data on missed care or mortality, allow us to examine these phenomena from a broader and less context specific base (Ausserhofer et

al. 2014). The scale and multi-site, multi-country design of studies such as Study IV, serve to increase the likelihood of external validity and reduce the risk of bias.

However, in order to explore the variables of interest, such as the relationship between nurse staffing and care left undone, the differences between countries are 'controlled for' by the country variable being treated in the regression analyses as a fixed or random effect. The goal in such analyses is to neutralise the differences that may relate to the country, in order to make the cases more closely equivalent and to isolate the relationship of interest. Regardless of *how* country effects are controlled for (e.g. whether fixed effect or random effect in regression models), multi-country studies necessarily reduce the contextual complexity associated with country differences, so that the relationship of primary interest, in this case staffing levels and patient outcomes, can be studied with less analytical 'noise'.

Yet even from my own limited experience of working with data from different countries, through being a member of the 15 country RN4Cast consortium, reveals that there are significant, albeit at times not clearly visible differences between countries in relation to some of the most fundamental aspects of the research. For example, what constitutes a hospital, a unit, a RN, a support worker? At times, members of the international team may have taken for granted that we had a shared language of health services research and nursing care. More exacting scrutiny revealed some significant differences not just in healthcare systems and hospitals, but in how nursing is organised and operationalised in each country.

Study III set out to consider more about the context of care left undone by examining it in a different country context – Sweden – and taking a descriptive look at the care context and how it compares to England. This examination allowed us to identify whether the relationships found between care left undone and staffing, in Study II in England, were also found in Sweden. Some of the 'other' factors that shape the context of nursing care provision in which these relationships exist also surfaced; factors which might otherwise have been hidden from view.

Staffing levels on Swedish wards varied according to the time of shift, with early/day shifts being better staffed than afternoons. RNs in England were more likely to report that they had a primarily direct care role (Ball et al. 2012) than were RNs surveyed in Sweden. Swedish nurses are less likely than English nurses to report that they often do the potentially 'transferrable' activities, or 'non-nursing' activities. This supports the picture gained from other research that RNs in Sweden are less likely to be in direct care roles than was even the case among RNs in Norway or Finland, other Nordic countries with several other similarities with Sweden (Lindqvist et al. 2014).

The findings reveal differences in nurse roles and deployment patterns between England and Sweden. For example RNs who report they often undertake potentially transferrable activities were more likely to report that they left care undone on their last shift. Added to this there is a more general difference in staffing levels between the two countries, with wards in Sweden

typically having better staffing levels of both RNs and of support workers; for example overall staffing on day shifts was an average of 2.9 patient per member of nursing staff, compared with 4.3 in England. These issues, taken in combination, may contribute to the difference observed in the marginal benefits of nursing support staff in Sweden, but not in England.

Despite the contextual differences, there was a notable degree of consistency in the relationship between RN staffing levels and care left undone in Sweden and England.

7.1.7 Care left undone is a predictor of case-mix adjusted hospital mortality

In a nine-country analysis of outcomes following common surgical procedures, care left undone was significantly associated with case-mix adjusted 30-day mortality. (Study IV)

In a large-scale multi-site multi-country study, we have demonstrated, for the first time, a significant association between care left undone and case-mix adjusted patient mortality. A 10% increase in the percent of 'care left undone' was associated with a 16% (OR 1.159 95% CI 1.004-1.026) increase in the likelihood of a patient dying within 30 days of admission following common surgical procedures.

The relationship between RN staffing and mortality has frequently been reported (Kane et al. 2007), and to a lesser degree so has the relationship between staffing and missed care (Jones, Hamilton, and Murry 2015; Cho et al. 2015; Kalisch, Tschannen, and Lee 2011). Fewer studies still have examined the potential consequences of care being left undone, such as the increased risk of patient falls (Kalisch, Tschannen, and Lee 2012). Research to date has not established an association with mortality.

Study IV built directly on the RN4Cast nine-country analysis that established relationships between RN staffing, nurse education levels, and hospital mortality, for patients following common surgery (Aiken et al. 2014). The analysis used the same dataset, based on the outcomes of 422,730 patients who underwent common surgeries in 300 European hospitals, and due to the consistency in both the data and the analytical and authorial team, the approaches to generating each of the measures from the data have been as closely replicated as possible. The only factor introduced to the analysis was care left undone.

The advantage of scale that a multi-site, multi-country design yields is the power needed to investigate differences between hospitals; power that has previously been lacking. But the result is that only limited account is taken of the within-hospital, within-country and between-country variation in any of the factors considered: staffing, care left undone, mortality. As Study III illustrated and other data has described, there is a significant degree of variation in what RNs do, and how this varies from place to place, as well as between the countries within the RN4Cast dataset (Aiken et al. 2012). Intra-hospital variation, found in both RN staffing and outcomes, may mask relationships between staffing and mortality when the relationships are examined at the hospital level (Van den Heede et al. 2008).

Nonetheless the findings – that care left undone contributes to an increased risk of patients dying following common surgery due to factors other than their condition – are evidently important from a patient safety perspective. Whilst there are ongoing debates about how best to define and measure them, errors in hospital remain a major cause of death since the Institute of Medicine highlighted it as a patient safety issue in their 1999 report 'To Err is Human' (Makary and Daniel 2016; Kohn, Corrigan, and Donaldson 2000; Shojania and Dixon-Woods 2016). Many patient safety initiatives have focussed on reducing errors of commission (Reason 1990). The relationship established in Study IV between care left undone and post-surgical mortality suggests that reducing errors of 'omission', such as care that is left undone, may be equally important (Kalisch, Landstrom, and Williams 2009).

As missed care acts as predictor of increased risk of patient death due to low staffing, there is potential for hospitals to use a measure of missed care as a 'leading indicator' rather than rely on measures of mortality, with all the inherent challenges of suitable patient-mix adjustment.

7.1.8 'Care left undone' is a mediator

Care left undone mediates the relationship between RN staffing and patient mortality; it presents a plausible causal pathway between RN staffing levels and patient mortality. (Study IV)

Whilst there have been many studies relating RN staffing to mortality rates, the inference that the relationship between nurse staffing and patient mortality may be causal, has been just that; an inference. Despite theories and speculation as to the nature of the association, possible mediating mechanisms have not previously been tested. Needleman's recent review and discussion paper raised the identification of explanatory mechanisms for the association as a research priority (Needleman 2016). Having established that variation in post-surgical mortality is related to missed nursing care in the first part of Study IV, we went on to examine whether missed nursing care mediates the relationship between nurse staffing and inpatient mortality.

The literature on patient safety has typically examined the direct effects between system-level structures and patient mortality, with little attention paid to potential mediating factors, such as missed nursing care. Mediation analysis has found that care left undone explains the association between nurse staffing and patient satisfaction with care (Bruyneel et al. 2015). However, research has not sought to investigate how missed nursing care might relate to patient mortality, and whether missed care is a mediator between RN staffing and patient mortality.

A PubMed search for original research articles published between Jan 1, 1985, and Sep 30, 2016, with the search terms (separately and in combination) "outcomes", "mortality", "mediation", and "indirect effect", found no papers published that have used mediation analysis to explain the association between nurse staffing and patient mortality.

The results of the analyses in Study IV are striking: missed nursing care is not only associated with patient mortality, but is also found to mediate the relationship between nurse staffing levels and patient mortality rates. Clearly the data and measurement caveats that apply to Study IV in general also apply to this finding: reliance on self-report data for staffing and care left undone, and that the study focusses on a single patient group.

The findings nonetheless offer support for a plausible casual pathway between nurse staffing and patient mortality, supporting the potential existence of a causal relationship between nurse staffing and outcomes; an interpretation of the relationship that has hitherto been contested (Griffiths et al. 2016).

There are multiple interpretations of the finding regarding care left undone as mediator between RN staffing and increased risk of mortality. What are the possible explanations for the relationship observed between care left undone and increased risk of death in hospital? Is it the absence of fundamental care - i.e. what the patient did not receive, but was needed for their wellbeing and recovery - that leaves patients more vulnerable to decline and death? Such an explanation would reinforce the importance of fundamental nursing care (Kitson et al. 2010) and its provision by RNs, to the well-being and recovery of medical and surgical patients.

Alternatively, it could be that nurses who repeatedly leave care undone due to lack of time, and lack of staffing, become habituated to it, and this is why care left undone is associated with higher risk of patient death. Does working in environments where incomplete care has become the 'norm' produce a nursing workforce that is no longer attuned to signs of deterioration? If this hypothesis were true, it is not the absence of care per se that creates the risk, but the presence of constant compromise that is the risk factor that puts patient lives in jeopardy.

Or perhaps both these interpretations place too much emphasis on the individual RN and what they do and how they respond to workload pressure, when there may be organisational level explanations for the relationships observed. What makes some hospitals staff their acute medical and surgical wards with half the level of RN staffing, compared to other hospitals? The differences in RN staffing levels appear not to reflect differences in patient dependency and requirement for nursing input (as seen in Study III). So we are left speculating on what this apparent mismatch may signify; are there fundamental flaws in the funding allocations that underlie the staffing differences observed? Do differences in the RN staffing levels observed reflect differences in the value placed on RN contribution? Or are low achieved RN staffing levels reflective not of differences in the scheduled staffing, but in the ability of hospitals to achieve the staffing levels planned, but to recruitment and retention problems?

The finding that care left undone mediates the relationship between RN staffing and mortality sparks as many questions as it answers.

7.2 METHODOLOGICAL ISSUES AND LIMITATIONS

There are numerous, significant challenges in researching the relationship between nurse staffing and patient mortality in hospitals which apply to the field in general but also to my research in particular. Key issues are: how we measure and isolate RN staffing, how we measure mortality that is potentially 'avoidable', how we take account of the complex and varied contexts, and how we capture potentially intermediate variables that may mediate the relationship, such as care left undone. The methodological challenges and how they relate to the research in this thesis are discussed in this section.

7.2.1 Isolating RN staffing as an input

Firstly, there is the challenge of defining and isolating the 'intervention' itself - RN staffing levels - from other attributes of the nursing workforce deployment such as the quantity of the contribution from nursing support staff or the educational levels of the RN workforce. The most immediate challenge is how best to quantify RN staffing levels. As with many methodological challenges, there are a number of separate but connected underlying measurement issues: the choice of measure (and unit of measurement associated with it), the source of data for the measure, and the computational method for deriving the measure from the data source.

The primary staffing measure used in this thesis, and in the RN4Cast study more generally, has been the RN: patient ratio of shifts. The advantage of this measure is that it can be easily captured through a survey of nurses. It requires collection of two pieces of information, the total number of RNs and the total number of patients; information that is available to nurses at every shift handover. In many studies, and in the case of RN4Cast, nurses are asked to report on the most recent shift they worked, so the recall period is kept to a minimum. It provides an empirical measure that can be readily understood, and measures the achieved deployed staffing, as opposed to the number of beds per post, or the level of staffing rostered (but not necessarily achieved, or the gap between planned and actual staffing (staffing deficit or 'short-staffing'). Using this measure allows the association between staffing and quality of care measures or outcomes to be examined. But beyond this, it also enables the RN: patient ratio associated with specific levels of quality or outcome attainment to be expressed. For example, in Study II a staffing level of six patients or fewer per RN was associated with a significantly reduced risk of care being left undone. It therefore has potential utility, depending on how it is used analytically and how it is reported, to inform policy and practice about RN staffing levels associated with good or poor outcomes.

Limitations however are that the ratio is derived from a survey, and is based on nurses' recall. It takes an unstructured sample of shifts to derive averages, and provides a one-off cross-sectional view of staffing without being able to describe changes in staffing over time, in relation to outcomes. On the other hand, the advantage for research is that it is a measure based on primary data-collection that can be applied in a consistent way in any hospital, in any

country (as it was in RN4Cast), without reliance on the availability of consistent administrative staffing data.

Nursing hours per patient day (NHPPD) is an alternative unit of measurement and one that can be readily derived from administrative data, if rostering and staffing data are routinely collected and held, for example in an electronic rostering or nurse staffing tool (Needleman et al. 2011; Duffield et al. 2011). The measure takes the total number of RN hours of care provided (based on how many RNs were on duty) and divides it by the number of patients within a 24 hour period. The significant advantage is that being based on routinely collected data, the measure is not restricted to a sample of shifts, but is potentially available for every day of the year, for as long a period as a study requires. This also means change of staffing over time can be examined, and it locates the timing of staffing inputs. From this there is the potential to establish ward staffing levels that preceded a defined patient outcome (such as death). However, when expressed as NHPPD – for the 24 hour period – this measure does not show difference in staffing levels at different times of the day or night, but effectively reports an averaged staffing level across 24 hour periods.

The NHPPD can be used to generate a measure of staffing deficit: for example comparing the achieved NHPPD relative to the planned NHPPD for each shift. Where the achieved level of staffing matches the planned level, the deficit would be zero. This can then be used to examine the impact of patient exposure to a specified level of deficit, on a measure of care quality or outcome. For example one RN too few on an eight hour shift, would be a deficit of eight hours. The advantage of this approach, is that where a systematic approach has been used to determine the nursing hours of care required, the 'planned staffing', it measures staffing deficit that is customised to each ward, based on the mix of patients and different level of dependency. For those hospitals that have a robust system in place to assess nursing workload and determine the staffing required for each shift, it effectively produces a measure of 'low-staffing' that is case-mix adjusted and specific to each ward, rather than relying on a post-hoc definition.

However, where the level of RN staffing planned has not been based on an empirical assessment of need, identifying the difference between planned and actual will not generate a consistent measure of 'deficit'. This was a criticism of a similar metric introduced in England in 2014, following the Francis inquiry: hospitals reported the average 'fill-rate' i.e. the proportion of shifts that had a full complement of staff. It relies on the 'planned' level being a reflection of the level required, for measures of the 'achieved' relative to 'planned' level to be meaningful.

An alternative measure that uses administrative data sources, where staffing data is not captured per shift or per 24-hours, can be computed based on the total number of posts relative to the number of beds (as used for the national data set in Study I). Most typically this is at the whole hospital level (or whole hospital group such as Trust in the case of England), as the measure is used where ward level data are not held, or are not accessible. An advantage of this measure is

that it can capture staffing for multiple staff groups – for example for medical, RNs and support staff in the national data set used in Study I – in the same format, regardless of different rostering systems used and the area covered by different staff groups (some ward based, some working across a specialty area).

However, several aspects of this measure represent a compromise to some degree. It is based on the number of substantive posts (the number of staff employed), which due to staff sickness, maternity leave and long-term absence etc., may differ from the number of staff working in the hospital. It does not take into account the presence of other staff working who are not directly employed by the hospital trust but are temporary staff employed by staffing agencies. In its crudest form (where services cannot be disaggregated) it is based on the total number of staff employed anywhere within the hospital, including outpatient services, day care, and rehabilitation. Added to this, some care providing organisations are administrative centres that may include services beyond general acute hospital care, such as community care or mental healthcare. The posts related to these services will not be contributing to the staffing per acute bed, but will nonetheless be counted in the numbers, producing some distortion of this as a measure of staffing that might relate to the outcomes of inpatients.

In order to isolate the potential impact of RN staffing from the contribution of others, requires the inclusion of data on staffing levels for other staff groups. Another challenge, hampering the research in general and Study I in particular, is the lack of available sources that enable consistent, detailed data on staffing levels of other staff groups within the multidisciplinary team, that go beyond this crude measure of posts to beds, to a more sophisticated measure of hours per patient day, for each group.

Although ward level data on staffing levels for support staff is more readily available than for medics, there are nonetheless issues to be considered in how the contribution of support workers is measured and considered. Whilst RN staffing levels and the proportion of nursing staff that are registered (defined as nursing skill-mix) may be closely connected operationally, a richer skill-mix does not equate to a larger overall 'dose' of RN care per patient. A measure of skill-mix (proportion of the nursing team that are RNs) thus does not provide an empiric measure of the 'size' of input from non-registered nursing support staff. A larger total nursing staff with a dilute skill-mix may nonetheless have higher levels of RN staffing than a neighbouring ward with a rich skill-mix but a relatively small team in total. For this reason, analyses that treat RN staffing levels and support worker staffing levels as separately defined measures permit the contribution of each to be described and interactions between the two to be studied (as in Study II and III).

7.2.2 Measuring 'avoidable' mortality

Study I and Study IV both included a measure of case-mix adjusted mortality as the outcome variable. Using mortality rates as an indicator of hospital performance requires the assiduous

application of risk adjustment in order to take account of the multiple and complex patient factors that contribute to risk of death. Whilst death may be a concrete and finite outcome, arriving at a measure of the rate of potentially avoidable deaths in hospitals is far more complex. The concept of 'standardising' or 'case-mix adjusting' mortality rates so that higher than expected rates of death can be identified, and a measure of potentially avoidable mortality derived continues to be an issue for debate (Shojania and Dixon-Woods 2016).

'Hospital care factors' that may potentially increase risk of patient death are many, varied, and complex. Inputs from a range of different staff, the decisions taken, treatment protocols, nutrition and hydration provided, medical care, drugs received, care environment and facilities available, are all conceivably factors that may influence the outcome of patients, and could confound or interact with the effects of the 'dose' of nurse staffing. As Griffiths notes, whilst high mortality rates may arise from deficiencies in nurse staffing, hospital mortality rates may not be sufficiently sensitive an indicator to detect the possible effect of RN staffing, since many other causes intervene (Griffiths et al., 2014).

Most deaths in hospital or following hospital care are caused by the patient's condition – the severity of their illness/presenting condition, and the range of other health issues present that impinge on their ability to recover. As a result, what is sometimes referred to as 'avoidable mortality' i.e. deaths that relate to hospital factors rather than patient risk factors, is relatively rare. For example, in the paper from Study I, we reported that in the national population of hospitals, the average mortality rates were: 32.8 deaths per 1,000 medical admissions and 7.9 deaths per 1,000 surgical admissions. Using Hogan's analysis, we would estimate that 95% of the variation between hospitals in these observed mortality rates is likely to be patient related (Hogan et al. 2012). Despite the challenges and the potential lack of sensitivity, Study I nonetheless detected an effect of medical and RN staffing on levels of case-mix adjusted hospital mortality derived from administrative data.

Whilst Study I examined hospital rates of case-mix adjusted mortality, in Study IV the mortality of a single patient group was examined: those undergoing common surgical procedures. Focusing on a narrow set of patients/conditions has been described as one way in which the sensitivity of risk-adjusted mortality measures can be increased, so that differences in hospital performance are more likely to be discerned (Shojania and Forster 2008). The reservation that applies however, is that the findings relate to a particular patient group and consequently we do not know the extent to which the findings may apply more generally.

7.2.3 Complexity of context

The complexity and variability of the context in which the relationships being examined exist, means there are a great number of potential confounders that need to be taken into account: hospital characteristics, ward/unit characteristics, the contribution of other members of the team, and the quality of the practice environment (for example leadership, team-working,

resources, professional autonomy). A quantitative observational design makes use of the 'natural experiment' that wide variation in the variables of interest offers, but it has well-known limitations, in particular that correlation between two observed variables does not prove causality. This weakness is to some degree ameliorated in studies that adopt a design where the sequence is established: that the input variable (RN staffing) precedes the outcome variable (mortality).

Observational studies also vary in terms of the depth (small number of sites in considerable detail) and breadth (large number of sites, but less detailed data about each). For example, the Needleman study (Needleman et al. 2011) was based on a single hospital but had staffing data on a large number of shifts. In contrast, multi-country multi-site studies such as RN4Cast have data from large numbers of hospitals, but many of the nursing characteristics (including RN staffing levels, care left undone, and the practice environment) are based on self-reports by nurses on relatively few shifts (Aiken et al. 2014). The advantage of a Needleman type design, is that much of the context is held static, whilst the relationship between RN staffing and outcomes are considered. However the RN4Cast study design, with its multiple sites and surveys of nurses to measure a wide range of workplace and workforce characteristics, allows for some of the variation in the context, whether at ward or hospital level, to be captured and considered.

7.2.4 Identifying mediators and moderators

Finally, a major challenge, and one that has probably received least attention in the field in general, is trying to determine the factors that sit between RN staffing at one end and patient outcomes at the other. Identifying the possible mediators and moderators of the relationship, and developing and testing a theoretical basis for the relationships seen between RN staffing and patient outcomes such as mortality. Identifying what it is about the presence of nurses and their actions that is associated with a positive impact on patients. Or conversely, as it is examined in this thesis, how absence of nursing care may contribute to adverse patient outcomes, and at the extremes, to the risk of avoidable patient death in hospital.

As noted in the background, the three main instruments used to measure missed care, care left undone, or rationed care, rely on self-report by nurses, using a variety of terms and units of measurement. But a challenge that applies to all these measures of missed care, is that to ascertain what is missed, we must first make a judgment about what could be missed: what activities constitute nursing care. This touches on a topic that has preoccupied nursing academics for decades: how do we define what nursing is? (Leininger and Gilead 1984; Bjorklund 2004). But rather than be drawn into this existential discussion, a starting point can be to consider the activities that comprise nursing, drawing on recent work by Kitson and colleagues who define the elements of 'fundamental nursing care'(Kitson et al. 2010; Kitson, Muntlin Athlin, and Conroy 2014). The challenge is to condense these broad domains into discrete care activities that can be ticked as done or undone, and that will apply in different

specialty settings, and different countries. A challenge that the instrument used in the RN4Cast study has sought to address.

We know that aside from the more esoteric discussion of what is nursing, there are more prosaic differences in terms of how nursing care is divided between members of the healthcare team: the role boundaries between medical staff, RNs and support workers for example. There may well be wide variation even within a country, or within a hospital, in the activities undertaken by support workers, which impacts on the activities considered to be within the RN role (Kessler et al. 2010). Variation in the role of the RN, as detected in Study II and III, creates a challenge in arriving at a universally applicable set of care activities that could be deemed 'RN work', and which form the basis of work that identifies what is not done.

That said, the instrument used to measure care left undone was developed from work in both the USA and Europe, so can hardly be considered as offering a parochial or country specific view of nursing activity. It has been through the use of this 'standard' instrument, within a questionnaire that examines many other aspects of nursing deployment and the practice environment within which care is delivered, that we have been able to explore variation in role and in care left undone, for example contrasting findings from Sweden and England. For all the challenges associated with standardising measures, whether in regard to mortality rates, measures of care left undone, or staffing levels, the key benefit is that it allows consistently defined terms to be studied, and the relationships between them to be explored.

7.3 IMPLICATIONS FOR FUTURE RESEARCH

The most fundamental challenges in researching the associations between nursing inputs and patient outcomes, remain the reliance on observational data and, in the case of the research in this thesis specifically, the use of cross-sectional data and self-report measures of staffing and missed care.

Nonetheless, within the strictures of a cross-sectional study, there is scope to refine the approach to enable greater confidence in the findings generated. That aspiration is the motivation for the research in this thesis. Using data from a multi-country cross-sectional study, I have re-examined some of the 'knowns' in this field, paying greater attention to 'other' variables (such as other staff, other nursing deployment factors, the context of care) that may contribute to the established associations between RN staffing and mortality, and between RN staffing and missed care. The research culminates in using the cross-sectional data from nine countries to address a new question: does missed care mediate the relationship observed between RN staffing and mortality?

In reviewing the results and assessing them in the light of their methodological limitations, it can be argued that the research in this thesis has taken some small, but nonetheless significant, incremental steps, to increase our knowledge and understanding of RN staffing, care left undone and mortality. But what does the research in this thesis, and my experience of

undertaking it, tell us about how things could be done better? What more needs doing to deepen our understanding of the impact of nurse staffing on patient mortality, and the role of nursing care that is left undone?

Firstly, much of the research focuses on efforts to identify staffing insufficiency and the consequences of too few nurses. 'Too few nurses' is however a broad concept yet its breadth is rarely spelt out. The terms and constructs encompassed by 'too few nurses' can become conflated. 'Short staffing' is different from 'low staffing', and 'nursing shortages' are different again. There is overlap between these issues: a constrained labour market may lead to high vacancy levels, which contributes to greater likelihood that individual wards are short-staffed, and shifts struggle to have a full complement of RNs on duty, without increasing reliance on temporary staff or making substitutions. Although related they are nonetheless distinct; each of these factors is influenced by multiple factors. But the underlying constructs can become blurred with one another. When we talk about a 'nurse shortage' are we referring to a lack of RNs in the labour market because the demand for them has outstripped supply, and workforce planning has not been adequate to deal with this? Or do we mean that we have hard to fill vacancies, as attracting and keeping RNs is a challenge, perhaps because of poor working conditions including heavy workloads? Or do we mean that typically there are too few RNs on duty to get the work required done? Or that the scheduled number of staff is adequate, but wards are frequently running below the planned level – 'short-staffed' – because the number employed is not sufficient to fill the rota?

The vast majority of the research to date focusses on the deployed RN staffing level, whether measured by ratios or NHPPD. A few studies examine exposure to 'short staffing' i.e. fewer staff than planned and deemed necessary based on a workload planning tool. But we are not seeing how the staffing situation we observe has come about: why were there less staff than needed, why does the ward in question repeatedly have lower levels of RN staffing than a ward of comparable size and case mix? Rather than looking at RN staffing levels in isolation, much might be gained from contextualising the research and considering the influence of workforce policy and labour market influences that have a bearing on the achieved RN staffing levels.

Research has repeatedly focussed on either day-time staffing, or staffing averaged across the 24 hour period. Study III revealed significant differences in staffing, and in care that is undone, by the time of shift. A more concerted focus on time of day may reveal interesting differences in the relationships between RN staffing, care left undone and patient risk of death, and how these differ between days and nights. Such insight may be valuable to inform practice, in terms of rostering staff to ensure sufficiency on hospital wards throughout the 24 hour period, rather than considering overall NHPPD across the 24 hour period.

Likewise, whilst there has been considerable research into nurse staffing and patient outcomes in acute hospitals, there has been very little of the same type of research in other settings. Perhaps it is time to apply the best of the research that has been developed over the past thirty

years to other contexts: to consider nurse staffing and patient outcomes outside of the general acute hospital.

Quantitative research inevitably requires a degree of simplification and standardisation of variables of interest, to make them discrete and consistent for entry into statistical models. During the course of my research I have become increasingly curious about other dimensions that are encompassed by 'RN staffing', beyond the overall RN: patient ratio, such as level of experience, grade, contractual status, and how these may contribute to the effectiveness of the nursing workforce.

Research findings and our interpretation of them depends on what we choose to study, which variables we decided to include. The choice of variables and factors included in the RN4Cast study design was based upon decades of work in this area. As described in the background, studies have built upon one another, often using similar instruments to measure the same things. The goal is to make use of the knowledge of what has been done before. Yet there is the danger that the knowledge that 'under-pins' an area of research could to some degree restrict it, rather than support it. In the RN4Cast study, we test the relationships amongst the regular 'players' in the work of nursing research. The presence of associations between these characteristics, does not mean that they are the only ones, or the most relevant ones. We do not know about the things we have not looked at or measured. Our thinking, in broadly replicating an approach to research, may be confining us to a particular viewpoint. There is a risk we may be pursuing the most readily measurable rather than the most relevant (Black 1996).

Few studies have adopted a different overall approach with the exception of the study by Needleman and colleagues (Needleman et al. 2011). In moving forward we need to think about how we can build on the significant body of knowledge that has been accumulated from the 'standard' cross-sectional design by using different approaches and techniques, taking inspiration and adopting the best of other designs to deal with some of the challenges identified.

Shekelle, in reviewing the evidence on nurse staffing and patient mortality concluded: *"To further advance this field, studies assessing an intentional change in nurse staffing ratios are needed. It may be impractical for such a study to be a randomized, controlled trial, but high-quality evidence could come from a time series analysis or a controlled before-and-after study, particularly if it included the necessary process variables to serve as a test of a conceptual framework for how increased staffing can influence outcomes."* p408 (Shekelle 2013b).

Future research could build on the Needleman approach, which achieves a closer alignment between the patient outcome and the nursing input, using a survival analysis of outcomes subsequent to nursing inputs (Needleman et al. 2011). But whilst that study was ground-breaking in moving our insight on RN staffing and patient outcomes forward, we need to be careful not to "throw the baby out with the bathwater". There are attributes of the practice environment, of the experience of delivering care, of work pressure, that nurses themselves

may be best placed to report on. Surveys allow a large number of people to contribute their views and report their observations and experience to researchers. One way forward could thus be to extend the Needleman design so that it is multi-site, and to augment it by including data derived from nurses, so that the influence of context could be more fully explored.

One of the key challenges for the research in this field has been to generate case-mix adjusted measures of mortality. There may also be scope to increase the sensitivity of case-mix adjusted mortality rates as an outcome if we not only adjusted for diagnoses and demographic factors, but also included a measure of the severity of patient illness at the time of admission – for example an early warning score system that includes measurement of key clinical indictors (Alam et al. 2014; Prytherch et al. 2010). Alternatively, research could focus more directly on potentially fatal clinical conditions, which are considered to be treatable. NHS England identifies a range of interventions which are likely to have a 'significant impact on reducing premature mortality' and includes the identification and treatment of sepsis for example (NHS-England 2013). The outcome for all sepsis patients could be examined in relation to the care input and context of care that had been experienced. For example, a measure of the hours per patient day for RNs, support workers and medical staff could be used, or could look at exposure to either 'short-fall' (deficit as defined per shift, per ward) or to a benchmark of low RN levels (such as more than eight patients per RN). (NICE 2014).

The research in this thesis has touched on differences between context of care, and the relevance of these to understanding how it is that nursing staff (both RNs and support staff) are deployed, what is done and undone, and the consequences of care provided, or care left undone. My exploration of how the context – 'other factors' – might impact on the relationships we seek to understand has been limited. It provides a hint that there may be much more going on than meets the quantitative researcher's eye. Entering different factors into a regression model to 'take account' of context offers a somewhat blinkered view. Perhaps it is time that we think the beyond quantitative research methods to examine RN staffing, care left undone and patient mortality. According to Black, the failure of health services research to use qualitative methods "has retarded the advance of medical knowledge and at times led to false trails being followed" p425 (Black 1996).

We may need to take a wider look at the research evidence base that might contribute to understanding the relationships between nurse staffing and care left undone, and between care left undone and adverse patient outcomes, to go beyond demonstrating associations to understanding why and how such relationships exist. For example a more deliberate attempt could be made to explore the evidence in a less restrictive fashion, seeking out research (of any sort) which informs the issues of how nurse staffing may relate to patient outcomes, regardless of whether RN staffing levels are specifically considered. Such an exploration may reveal evidence that helps explain how care is left undone, and factors that potentially exacerbate or mitigate the effect of workload pressure which lead to care being left undone. We could also begin to seek answers to some of the questions that the research in this thesis has raised about

how it is that care left undone impacts on patient mortality. For example to examine the extent to which absence of fundamental care, lack of surveillance or nurses habituation may offer some explanation for the association between missed care and increased risk of patient death.

The use of electronic data capture for some aspects of nursing care planning and delivery offers an alternative approach that allows missed care to be defined in relation to the care that was expected/planned, and relative to the activities done. For example a study currently underway uses routinely captured data from an electronic system which records vital signs observations to generate a measure of delayed or missed care (Griffiths et al. 2015). This research builds on the idea that missed surveillance/monitoring may be a critical factor that leads to patient deterioration and risk of avoidable death.

We could also turn attention to explore in greater depth nurses' own experience and views of leaving necessary care undone. Study II and Study II revealed differences in the frequency that different activities were missed, and differences in the effect of staffing on the likelihood that care of different sorts was missed. Further research might explore the extent to which these differences relate to RNs efforts to prioritise care. In the face of time pressures, how do nurses decide what should be done and what to leave undone? A study that shadowed nurses during their shift, and interviewed staff at the end of shifts would potentially reveal something more about the RN experience of delivering care and leaving some care undone, and how this might vary between individuals depending on their experience or context.

But further than this, qualitative research through interviews, diaries, or focus groups, might be able explore the professional judgement and experience of clinicians such as RNs and doctors as to the circumstances that put patients at risk of fatal deterioration. Research could focus on trying to identify the activities, decisions, observations and coordination that is critical to keeping patients safe from avoidable harm, or that when absent, heighten their chances of dying in hospital unnecessarily.

7.4 IMPLICATIONS FOR POLICY AND PRACTICE

Every day within a hospital decisions are taken about how to staff a ward, how many RNs are going to be needed for a day shift, how many on a night, and judgments made about whether there are sufficient staff to provide care safely and to a high quality: to 'do the patient no harm'. Yet inquiries into hospital malpractice and patient safety breaches, and the conclusions of those charged with developing guidelines for 'safe nurse staffing', pointed to deficits in the evidence base and its application to practice.

My aim in conducting this research has been to understand the relationships, and theoretical explanations for them, between RN staffing, care left undone, and mortality, in order to inform policy and practice. Its purpose, as opposed to its aim, has thus always been firmly rooted in the attempt to address some of the limitations in the evidence (identifying uncertainties, lack

of specificity, utility, and the inference of causality), so that the findings from research can be more easily translated into policy and practice.

The descriptive elements of these studies have been important in documenting the level of care compromise that had previously not been described – in either England or Sweden. The findings from Study II were published in England at a time of increasing concern regarding patient safety and the adequacy of nurse staffing levels in NHS hospitals. A public inquiry recommended that national guidance on 'safe nurse staffing' levels be provided. However, the goal of determining 'safe staffing' levels to recommend in guidance revealed the limitations of the evidence base. Limitations which the studies in this thesis have gone some small way to addressing: by increasing confidence in the conclusions emerging from the evidence base, making explicit the levels of RN staffing associated with care that is less incomplete, and demonstrating a viable causal pathway connecting nurse staffing to patient outcomes via missed nursing care. For health service managers and policy makers, the research findings provide an account of the RN staffing levels associated with an increased risk of missed care. The analysis has allowed the associations at different levels of RN and support nurse staffing to be expressed.

The findings have implications for healthcare systems: for hospital managers responsible for making decisions about workforce size and configuration and deployment patterns, for regulators charged with monitoring patient safety, and for policy makers with responsibility for developing guidance on minimizing risk in hospitals to keep patients safe from avoidable harm. Appropriate workforce policies and practices are needed to minimize the risk of care being missed and of patients dying from factors that are within the health system's control, such as safe nurse staffing.

The findings point to the potential utility of 'care left undone' as a leading indicator of inadequate staffing levels. An idea that has subsequently been taken up by the health department in Ireland, who are trialling the use of 'Care Left Undone Events' to monitor the adequacy of RN to meet patient needs, shift by shift (DoH-Ireland 2016).

Many a good research study, and many a fine dissertation, have sat gathering dust on a book shelf, or unvisited in the archives, rather than being translated into practice. The potential relevance to policy and practice, does not mean that research findings are destined to be implemented. Implementation requires aspiration, intention and agency: to want to do it, to have plan to do it, and to have the means of doing it. Issues that are beyond this research, but that are key to yielding a benefit from it.

8 CONCLUSION

The research in this thesis has built on decades of research that precedes it and established with more certitude that RN staffing levels are related to patients' risks of dying in hospital, due to factors other than their condition. Taking account of a large number of potential confounding factors, the research has established that where RN staffing levels are lower, there is a higher risk that necessary nursing care is left undone. Despite the fact that the context in which nursing care is delivered on hospital wards varies considerably between countries or at different times of the day, the relationship between RN staffing and care being left undone is found consistently in multiple settings. In a large-scale study, with 300 hospitals from across nine different European countries, I, with my colleagues, have again established the relationship between RN staffing and hospital related mortality, observed so frequently before. For the first time the relationship between missed care and mortality has been examined using a sufficiently powered design, and the results found to be unequivocal: higher levels of missed care is associated with an increased risk of patient death in hospital following general surgery. When both staffing and missed care are looked at together, we find the effect of RN staffing on mortality is lessened. Missed care partially mediates the relationship between RN staffing and patient risk of death in hospital.

These findings support the imputation of causality in the relationship between RN staffing and patient mortality. Taking this research together with the studies that have preceded it, particularly those using a different methodology, we can conclude with some confidence that higher RN staffing levels reduce the risk of fatal harm to patients, and that this correlation is likely to be causal.

This evidence now needs to be used to inform health service practice and the workforce policies that guide it, so that the antecedents to patient safety are in place: sufficient RNs on duty to meet patient needs without leaving care undone, and reduce the risk of patients dying as a result of not getting the care they needed. Without a willingness to use evidence and translate it into practice, academic research remains just that, 'academic'.

Who is culpable for the death of patients that arise due to lack of nursing care provided, and too few nurses on duty with the right skills to keep patients safe? The individual nurses on that ward? Or those that have made the decisions about how to staff the service. And surely John Stuart Mill's words on inaction apply equally to policy makers as they do to hospitals, and nursing care-givers: "A person may cause evil to others not only by his actions but by his inaction, and in either case he is justly accountable to them for the injury". Applying this philosophy, failure to act, to apply the evidence on RN staffing to practice and policy, would render policy makers and those with responsibility for health service workforce also accountable for injury that arises due to insufficient nurse staffing.

The evidence base is not perfect – it never is. Knowledge and science are constantly evolving. Each study, including those in this thesis, has limitations. Nonetheless, the conclusions arrived at when placing this research together with the wider body of evidence, are well supported.

That does not mean that there is nothing further to do. We need to find out more about the intermediary factor, care left undone, to better understand how the configuration of the healthcare workforce contributes to the risk of patient death in hospital. We have perhaps gone as far we can within the limits of the cross-sectional design in researching the relationship between RN staffing and patient mortality in acute general hospitals. To go further in this area requires new thinking, new approaches, different datasets, and different questions. In this way we can starting unpicking the detail of how nurse staffing levels impact on outcomes, and the reasons that patients continue to die in hospital when they need not, 150 years after Florence Nightingale espoused her first principle: the hospital shall do the sick no harm.

Will having more or better research help address the problems of RN staffing levels? Will creating nurse staffing guidelines based on the evidence base make a difference? It is hoped so. But arguably the biggest problem in achieving safe nurse staffing levels is not that we do not know the numbers needed, or that we lack the means of determining appropriate RN staffing levels (Ball and Catton 2011). Perhaps the most significant hurdle is a failure to fully comprehend the value of skilled nursing care provided by RNs. So that the workforce is not planned with diligence, staffing levels are not maintained in the way they are in other safety-critical services. Which may be why missed care is an important measure; a measure that is reflective of RN staffing insufficiency and predictive of higher patient risk. Through examining the negative space that is 'care left undone', by looking at the consequences of absence, we are perhaps better able to see the value of presence. The value of nursing care delivered by RNs to the outcomes of patients in hospitals.

9 EPILOGUE

It's 1988. I'm sitting there, 23 years old, with my bleached blonde hair-do, my leggings and belted baggy shirt, waiting to meet with the boss's boss: Head of Department, Professor David Canter. The cheek of it really. Thanks to my good friend and inspirational academic Dr Margaret Wilson, I'd managed to get a few weeks doing data entry in the Psychology department at my alma mater, University of Surrey. I'd jumped at the chance, the small window of opportunity into the world I found so alluring: research. Permission to ask questions and try and answer them. After some weeks at Surrey, I'd been upgraded from being a temp into a newly formed position 'research administrator'. A 'niche' post that allowed me, with my somewhat unremarkable 2:2 BSc in Nursing, to be given a chance to assist in the research of a thriving academic department. My credentials were lacking but they gave me a chance; for that, I'm very grateful.

I loved it. It was like a research apprenticeship: data entry, content analysis, coding, learning SPSS syntax, collecting output from the library, proof reading articles, helping organise events and sitting in on research group meetings, methods lectures, and conference sessions. An interesting, intelligent, inspiring mix of people applying psychology to research. And I got to see it and hear it and even be a part of it: taking it all in. So here I was, waiting to speak to Prof Canter about possible next steps for me in research. Should I think about doing a PhD perhaps? 'The most important thing, is that you do something that truly interests you: that excites you, puts a fire in your belly' he said. He was right. Yet I didn't have a burning desire to find out more about a particular topic back then. But I did love research. So I kept doing it.

I went on to work at the Institute for Employment Studies, where my nursing degree was an asset for research on healthcare workforce. Working with great people who were generous with their time and encouragement and helped me develop as a researcher, to learn new skills, run projects. Dr Dick Waite who taught me the benefit of well-kept references, and who, despite his far superior intellect, made me feel like I had something useful to offer to our research study on nursing roles in high-tech areas of care. And Prof Jim Buchan, who gave me the chance to write 'Caring Costs' with him and go on to work on many more healthcare workforce studies since. And who has continued to offer me mentorship over the years – thank you. Great colleagues, but none so special in every way as Geoff Pike, destined to be my partner for life.

Fast forward ten years, to 1999, and there I am, in a swish restaurant in London, meeting Prof Anne Marie Rafferty to talk about the possibility of doing a PhD. I had been hugely inspired by Celia Davies book: Gender and the Professional Predicament of Nursing. (Read it, it doesn't age). Perhaps I could find something in there to follow: the apparent

discrepancy between the level regulation of nurses compared with medics maybe? But Holly was three, and Amy five. And life working with Geoff in our own workforce research consultancy was sweet. Did I want to take four years out to work on my own? Perhaps I should wait until the children were older and I had more time I reasoned. But life doesn't pause.

More research, surveys of nurses about their working lives for the Royal College of Nursing (RCN), working with the inspirational Pippa Gough on staff recruitment and retention and then thanks to Howard Catton a change of tack: a part-time post as policy adviser at the RCN. My job to use evidence to support the development of guidance on safe staffing for nurses in the UK. Working on the other side of the table made me see the importance of research that can inform policy and practice. And meanwhile thanks to Prof Anne Marie, who never gave up on me, and new colleague Prof Peter Griffiths, I spent the rest of the week working as a visiting fellow at King's College London, leading the nurse survey for the RN4Cast study. An opportunity to be part of large consortium of researchers: 30+ people meeting several times a year to think about the relationship between nursing workforce and patient outcomes, and how understanding it better could inform nursing workforce planning. Just the sorts of issues that I'd spent years touching on. Led by Prof Linda Aiken and Prof Walter Sermeus – this study was going to be big. But still I'm there with the name badge without the 'Dr' prefix. Hoping that my experience would be what counted, rather than my lack of title.

And I had no real desire to be an academic. I was doing the research I loved, and had a growing interest in its interface with policy. Why change? Because opportunity beckoned: to work at the National Nursing Research Unit, as deputy director. I could carry on working with King's on the RN4Cast study but now as part of the home-team, not a visitor. Again, they took a chance on me, considered my experience sufficient to compensate for the lack of a PhD. Working with director Prof Jill Maben was the most fabulous job: to be doing research with the explicit purpose of informing policy on nursing. What could be better? And to be in a unit with such extremely clever, yet modest and warm people, from such a mix of backgrounds. I felt honoured and privileged, and will always be very grateful for that job and the opportunities it gave me to engage in national policy whilst generating research. I arrived puzzled by the world of Academia, and unsure I would fit in. But thanks to the incomparable leadership and management of Prof Jill Maben, I found my feet, and saw the value of an academic life to research on nursing workforce policy.

But still, no PhD. At the age of 48, surely that ship had sailed? Or had it? Prof Peter Griffiths at the final RN4Cast consortium dinner in Brussels, made me think again: "I really think you should do a PhD. Not because you need it as such, although there may be things you'd gain from it, but because why shouldn't you? Others with PhD get the cudos associated with it, why shouldn't you too? Don't you deserve the same recognition and opportunities?' And Carol and Rikard, our Swedish RN4Cast colleagues, surprised

to discover this research cuckoo in the nest, spontaneously encouraged me to do it; apply to do a PhD with them, at the Karolinska Institutet. And thanks to the marvellous Swedish system, that recognises the value of doctoral students (and do not charge fees) and the benefit of PhD by publication, it perfectly fitted with my needs. Four years where I could carve and refine the area of research myself, but would have support and team-members to work with to produce published outputs. It's a great system and I am very grateful for it.

Four years later, 2017, and thanks to my supervisors, and all those people along the way in my career, and in my life, I'm completing my PhD thesis on nurse staffing and care left undone. A topic I care deeply about. And thanks to the quality of my supervision, to my lead supervisor Prof Carol Tishelman in particular, the experience of doing a PhD has far exceeded my expectations. It's given me a chance to take the many threads from my research career, and adding some new ones, weave them into a picture. I didn't know I had it in me, but I am so very grateful that with your patience, scrutiny, challenge, encouragement to be curious, to think again, to see it differently, to be more exact, I have completed this work.

To each person who ever sowed the seed of the idea, and to everyone that has helped me make 'doing a PhD' not just feasible thing, but a meaningful endeavour, 'Thank You'. This has been the best experience of my life as a researcher and I am truly grateful for it.

And to anyone reading this who has this journey ahead of them: choose something that matters to you, be curious about it and surround yourself with people who believe in you.

10 ACKNOWLEDGEMENTS

I should like to acknowledge and thank:

- All the RNs who took the time to complete the survey – without you this research truly would not have been possible. I hope the findings may be put to good use where you work.
- The European Union's Seventh Framework Programme (FP7/2007–2013) for their perspicacity in funding research on nursing workforce in general, and the RN4Cast study in particular.
- Vårdförbundet (Swedish Association of Health Professionals), The Swedish Council for Working Life and Social Research (FAS/FORTE), and Karolinska Institutet Strategic Research Programme in Care Sciences (SFO-V) for for their funding and support of the study in Sweden.
- Prof Linda Aiken and Prof Walter Sermeus for their steady and effective leadership of the RN4Cast study.
- RN4Cast consortium members – a privilege to have worked alongside you.
- Luk Bruyneel in particular for sterling work on mediation and as joint first author on the fourth paper.
- Caroline Nicholson for your mentorship over the years at the times I most needed it.
- University of Southampton and King's College London for supporting me throughout the PhD.
- Colleagues at the Innovative Care research group at the Medical Management Centre, LIME, KI for your engagement, support and friendship, particularly Lisa Smeds-Alenius, Mia von Knorring, Maria Lindberg-Reinius, and Sara Runesdotter.
- Ingrid Smedberg, Petra Åhreus Bäckström, and Sophia Savage at LIME, KI for making the practical problems of studying from a distance disappear.
- And lastly, my supervisors Carol Tishelman, Peter Griffiths, and Rikard Lindqvist, for encouraging me, challenging me, and keeping me focussed.

REFERENCES

Abbott, A. 2016. "Culture of silence and nonchalance protected disgraced trachea surgeon (updated)." *Nature News*. Nature Publishing Group.

Aiken, Linda H, Sean P Clarke, and Douglas M Sloane. 2002. 'Hospital staffing, organization, and quality of care: cross-national findings', *Nursing Outlook*, 50: 187-94.

Aiken, Linda H, Sean P Clarke, Douglas M Sloane, Julie A Sochalski, Reinhard Busse, Heather Clarke, Phyllis Giovannetti, Jennifer Hunt, Anne Marie Rafferty, and Judith Shamian. 2001. 'Nurses' reports on hospital care in five countries', *Health Affairs*, 20: 43-53.

Aiken, Linda H, Sean P Clarke, Douglas M Sloane, Julie Sochalski, and Jeffrey H Silber. 2002. 'Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction', *JAMA*, 288: 1987-93.

Aiken, Linda H, and Patricia A Patrician. 2000. 'Measuring organizational traits of hospitals: the Revised Nursing Work Index', *Nursing Research*, 49: 146-53.

Aiken, Linda H, Walter Sermeus, Koen Van den Heede, Douglas M Sloane, Reinhard Busse, Martin McKee, Luk Bruyneel, Anne Marie Rafferty, Peter Griffiths, and Maria Teresa Moreno-Casbas. 2012. 'Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States', *BMJ*, 344: e1717.

Aiken, Linda H, Douglas Sloane, Peter Griffiths, Anne Marie Rafferty, Luk Bruyneel, Matthew McHugh, Claudia B Maier, Teresa Moreno-Casbas, Jane E Ball, and Dietmar Ausserhofer. 2016. 'Nursing skill mix in European hospitals: cross-sectional study of the association with mortality, patient ratings, and quality of care', *BMJ Quality & Safety*: bmjqs-2016-005567.

Aiken, Linda H, Douglas M Sloane, Luk Bruyneel, Koen Van den Heede, Peter Griffiths, Reinhard Busse, Marianna Diomidous, Juha Kinnunen, Maria Kózka, and Emmanuel Lesaffre. 2014. 'Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study', *The Lancet*, 383: 1824-30.

Aiken, Linda H, Douglas M Sloane, Jeannie P Cimiotti, Sean P Clarke, Linda Flynn, Jean Ann Seago, Joanne Spetz, and Herbert L Smith. 2010. 'Implications of the California nurse staffing mandate for other states', *Health Services Research*, 45: 904-21.

Aiken, Linda H, Herbert L Smith, and Eileen T Lake. 1994. 'Lower Medicare mortality among a set of hospitals known for good nursing care', *Medical Care*: 771-87.

Alam, N, EL Hobbelink, AJ Van Tienhoven, PM Van de Ven, EP Jansma, and PWB Nanayakkara. 2014. 'The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review', *Resuscitation*, 85: 587-94.

Andersson, Åsa, Catharina Frank, AML Willman, P-O Sandman, and Görel Hansebo. 2015. 'Adverse events in nursing: A retrospective study of reports of patient and relative experiences', *International Nursing Review*, 62: 377-85.

Audit-Commssion. 2010. "Making the Most of NHS Frontline Staff." In. London: Audit Commission.

Ausserhofer, D., B. Zander, R. Busse, M. Schubert, S. De Geest, A. M. Rafferty, J. Ball, A. Scott, J. Kinnunen, M. Heinen, I. S. Sjetne, T. Moreno-Casbas, M. Kozka, R. Lindqvist, M. Diomidous, L. Bruyneel, W. Sermeus, L. H. Aiken, R. Schwendimann, and Rn Cast consortium. 2014. 'Prevalence, patterns and predictors of nursing care left undone in European hospitals: results from the multicountry cross-sectional RN4CAST study', *BMJ Quality & Safety*, 23: 126-35.

Aylin, Paul, Alex Bottle, and Azeem Majeed. 2007. 'Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models', *BMJ*, 334: 1044.

Ball, J., G. Pike, P Griffiths, A. M. Rafferty, and T. Murrells. 2012. 'RN4Cast Nurse Survey in England'.King's College London. National Nursing Research Unit.

Ball, Jane, and Howard Catton. 2011. 'Planning nurse staffing: are we willing and able?', *Journal of Research in Nursing*: 1744987111422425.

Birkmeyer, John D, Justin B Dimick, and Nancy JO Birkmeyer. 2004. 'Measuring the quality of surgical care: structure, process, or outcomes?', *Journal of the American College of Surgeons*, 198: 626-32.

Bjorklund, Pamela. 2004. 'Invisibility, moral knowledge and nursing work in the writings of Joan Liaschenko and Patricia Rodney', *Nursing Ethics*, 11: 110-21.

Black, Nick. 1996. 'Why we need observational studies to evaluate the effectiveness of health care', *BMJ: British Medical Journal*, 312: 1215.

Blegen, M. A., C. Goode, J. Spetz, T. Vaughn, and S. H. Park. 2011. 'Nurse Staffing effects on patient outcomes: safety-net and non-safety net hospitals', *Medical Care*, 49: 406-14.

Bond, CA, Cynthia L Raehl, Michael E Pitterle, and Todd Franke. 1999. 'Health care professional staffing, hospital characteristics, and hospital mortality rates', *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 19: 130-38.

Bowers, Len. 1989. 'The significance of primary nursing', *Journal of Advanced Nursing*, 14: 13-19.

Bowling, Ann. 2014. *Research methods in health: investigating health and health services* (McGraw-Hill Education (UK)).

Bruyneel, Luk, Baoyue Li, Dietmar Ausserhofer, Emmanuel Lesaffre, Irina Dumitrescu, Herbert L Smith, Douglas M Sloane, Linda H Aiken, and Walter Sermeus. 2015. 'Organization of hospital nursing, provision of nursing care, and patient experiences with care in Europe', *Medical Care Research and Review*, 72: 643-64.

Buchan, J. , and J. Ball. 1991. 'Caring Costs. Nursing costs and benefits: a review for the Royal College of Nursing '. Institute of Employment Studies. Brighton.

Buchan, James. 2005. 'A certain ratio? The policy implications of minimum staffing ratios in nursing', *Journal of Health Services Research & Policy*, 10: 239-44.

Campbell, Michael J, Richard M Jacques, James Fotheringham, Ravi Maheswaran, and Jon Nicholl. 2012. 'Developing a summary hospital mortality index: retrospective analysis in English hospitals over five years', *BMJ*, 344.

Cayton, Harry. . 2012. '"Mandating staffing levels is not the answer to reducing poor care'." In *Health Service Journal*.

Charlson, Mary E, Peter Pompei, Kathy L Ales, and C Ronald MacKenzie. 1987. 'A new method of classifying prognostic comorbidity in longitudinal studies: development and validation', *Journal of Chronic Diseases*, 40: 373-83.

Cho, S-H, Y-S Kim, KN Yeon, S-J You, and ID Lee. 2015. 'Effects of increasing nurse staffing on missed nursing care', *International Nursing Review*, 62: 267-74.

Cohen, Mark E, Clifford Y Ko, Karl Y Bilimoria, Lynn Zhou, Kristopher Huffman, Xue Wang, Yaoming Liu, Kari Kraemer, Xiangju Meng, and Ryan Merkow. 2013. 'Optimizing ACS NSQIP modeling for evaluation of surgical quality and risk: patient risk adjustment, procedure mix adjustment, shrinkage adjustment, and surgical focus', *Journal of the American College of Surgeons*, 217: 336-46. e1.

Dall, Timothy M, Yaozhu J Chen, Rita Furst Seifert, Peggy J Maddox, and Paul F Hogan. 2009. 'The economic value of professional nursing', *Medical Care*, 47: 97-104.

Despins, Laurel A, Jill Scott-Cawiezell, and Jeffrey N Rouder. 2010. 'Detection of patient risk by nurses: a theoretical framework', *Journal of Advanced Nursing*, 66: 465-74.

Dimick, Justin B, H Gilbert Welch, and John D Birkmeyer. 2004. 'Surgical mortality as an indicator of hospital quality: the problem with small sample size', *JAMA*, 292: 847-51.

DoH-Ireland. 2016. "Interim Report of the Taskforce on Staffing and Skill Mix for Nursing" Dublin: Department of Health Ireland

Donabedian, Avedis. 1988. 'The quality of care: how can it be assessed?', *JAMA*, 260: 1743-48.

Duffield, Christine, Donna Diers, Linda O'Brien-Pallas, Chris Aisbett, Michael Roche, Madeleine King, and Kate Aisbett. 2011. 'Nursing staffing, nursing workload, the work environment and patient outcomes', *Applied Nursing Research*, 24: 244-55.

Duffield, Christine M, Michael A Roche, Caroline Homer, James Buchan, and Sofia Dimitrelis. 2014. 'A comparative review of nurse turnover rates and costs across countries', *Journal of Avanced Nursing*, 70: 2703-12.

Estabrooks, Carole A, William K Midodzi, Greta G Cummings, Kathryn L Ricker, and Phyllis Giovannetti. 2005. 'The impact of hospital nursing characteristics on 30-day mortality', *Nursing Research*, 54: 74-84.

Feo, Rebecca, and Alison Kitson. 2016. 'Promoting patient-centred fundamental care in acute healthcare systems', *International Journal of Nursing Studies*, 57: 1-11.

Francis, R. 2010. 'Robert Francis inquiry report into mid-staffordshire NHS foundation trust', *The Stationery Office: London*.

Gerdtz, MF, and S Nelson. 2007. '5–20: a model of minimum nurse-to-patient ratios in Victoria, Australia', *Journal of Nursing Management*, 15: 64-71.

Ghaferi, Amir A, John D Birkmeyer, and Justin B Dimick. 2009. 'Variation in hospital mortality associated with inpatient surgery', *New England Journal of Medicine*, 361: 1368-75.

Goranitis, Ilias, and Alastair Fischer. 2015. 'Measuring the contributions of prevention and treatment of illness to changing life expectancy with avoidable mortality as the outcome indicator: how to deal with new diseases', *The Lancet*, 386: S38.

Griffiths, P., J. Ball, K. Bloor, J. Briggs, C. Dall'Ora, D. Prytherch, G. Smith, P. Meredith, N. Sinden, D. Böhning, A. Maruotti, A. De Iongh, P. Schmidt, and A. Recio-Saucedo. 2015. 'Nurse staffing levels, missed vital signs observations and mortality in hospital wards: modelling the consequences and costs of variations in nurse staffing and skill mix (Research study outline)', Accessed 6th March http://www.southampton.ac.uk/healthsciences/research/projects/nurse-staffing-levels-missed-vital-signs.page.

Griffiths, Peter. 2012. 'Inaugural lecture: Doing Away With Doctors? Workforce research and the future of nursing', *Working Papers in Health Sciences*, 1: 1-11.

Griffiths, Peter, Jane Ball, Jonathan Drennan, Chiara Dall'Ora, Jeremy Jones, Antonello Maruotti, Catherine Pope, Alejandra Recio Saucedo, and Michael Simon. 2016. 'Nurse staffing and patient outcomes: Strengths and limitations of the evidence to inform policy and practice. A review and discussion paper based on evidence reviewed for the National Institute for Health and Care Excellence Safe Staffing guideline development', *International Journal of Nursing Studies*.

Griffiths, Peter, Jane Ball, Jonathan Drennan, Liz James, Jeremy Jones, Alejandra Recio, and Michael Simon. 2014. 'The association between patient safety outcomes and nurse/healthcare assistant skill mix and staffing levels and factors that may influence staffing requirements'. Reivew for NICE. University of Southampton.

Griffiths, Peter, Simon Jones, and Alex Bottle. 2013. 'Is "failure to rescue" derived from administrative data in England a nurse sensitive patient safety indicator for surgical care? Observational study', *International Journal of Nursing Studies*, 50: 292-300.

Hall, Celia. 2004. 'Young nurses 'too posh to wash'', *Daily Telegaph,* 11 May 20014.

Hartz, Arthur J, Henry Krakauer, Evelyn M Kuhn, Mark Young, Steven J Jacobsen, Greer Gay, Larry Muenz, Myron Katzoff, R Clifton Bailey, and Alfred A Rimm. 1989. 'Hospital characteristics and mortality rates', *New England Journal of Medicine*, 321: 1720-25.

Healthcare-Commission. 2009. 'Investigation into Mid Staffordshire NHS Foundation Trust'. London.

Heinen, M. M., T. van Achterberg, R. Schwendimann, B. Zander, A. Matthews, M. Kozka, A. Ensio, I. S. Sjetne, T. Moreno Casbas, J. Ball, and L. Schoonhoven. 2013. 'Nurses' intention to leave their profession: a cross sectional observational study in 10 European countries', *International Journal of Nursing Studiese*, 50: 174-84.

Hogan, Helen, Frances Healey, Graham Neale, Richard Thomson, Charles Vincent, and Nick Black. 2012. 'Preventable deaths due to problems in care in English acute

hospitals: a retrospective case record review study', *BMJ Quality & Safety*: bmjqs-2012-001159.

Holland, Paul W. 1986. 'Statistics and causal inference', *Journal of the American statistical Association*, 81: 945-60.

Jarman, Brian, Simon Gault, Bernadette Alves, Amy Hider, Susan Dolan, Adrian Cook, Brian Hurwitz, and Lisa I Iezzoni. 1999. 'Explaining differences in English hospital death rates using routinely collected data', *BMJ*, 318: 1515-20.

Jones, Terry L, Patti Hamilton, and Nicole Murry. 2015. 'Unfinished nursing care, missed care, and implicitly rationed care: State of the science review', *International Journal of Nursing Studies*, 52: 1121-37.

Kalisch, B. J., D. Tschannen, and K. H. Lee. 2011. 'Do staffing levels predict missed nursing care?', *International Journal for Quality in Health Care*, 23: 302-08.

Kalisch, Beatrice J. 2006. 'Missed nursing care: a qualitative study', *Journal of Nursing Care Quality*, 21: 306-13.

Kalisch, Beatrice J, Gay Landstrom, and Reg Arthur Williams. 2009. 'Missed nursing care: errors of omission', *Nursing Outlook*, 57: 3-9.

Kalisch, Beatrice J, Margaret McLaughlin, and Beverly Waller Dabney. 2012. 'Patient perceptions of missed nursing care', *The Joint Commission Journal on Quality and Patient Safety*, 38: 161-67.

Kalisch, Beatrice J, Dana Tschannen, and Kyung Hee Lee. 2012. 'Missed nursing care, staffing, and patient falls', *Journal of Nursing Care Quality*, 27: 6-12.

Kalisch, Beatrice J, and Reg Arthur Williams. 2009. 'Development and psychometric testing of a tool to measure missed nursing care', *Journal of Nursing Administration*, 39: 211-19.

Kane, Robert L, Tatyana A Shamliyan, Christine Mueller, Sue Duval, and Timothy J Wilt. 2007. 'The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis', *Medical Care*, 45: 1195-204.

Keogh. 2013. "Review into the quality of care and treatment provided by 14 hospital trusts in England." London.

Kessler, Ian, Paul Heron, Sue Dopson, Helen Magee, Danielle Swain, and Janet Askham. 2010. 'The nature and consequences of support workers in a hospital setting', *Final report. NIHR Service Delivery and Organisation programme*.

Kitson, Alison, Tiffany Conroy, Yvonne Wengstrom, Joanne Profetto-McGrath, and Suzi Robertson-Malt. 2010. 'Scholarly paper: Defining the fundamentals of care', *International Journal of Nursing Practice*, 16: 423-34.

Kitson, Alison L, Åsa Muntlin Athlin, and Tiffany Conroy. 2014. 'Anything but basic: nursing's challenge in meeting patients' fundamental care needs', *Journal of Nursing Scholarship*, 46: 331-39.

Kohn, Linda T, Janet M Corrigan, and Molla S Donaldson. 2000. *To err is human:: building a Safer Health System.* National Academies Press.

Kossarova, Lucia, Walter Holland, Ellen Nolte, and Martin McKee. 2009. 'Measuring 'avoidable'mortality: methodological note'. Brussels, Belgium: Directorate-General "Employment. Social Affairs and Equal Opportunities.

Kramer, Marlene, and Claudia Schmalenberg. 1988. 'Magnet Hospitals Part II: Institutions of Excellence', *Journal of Nursing Administration*, 18: 11-19.

Kutney-Lee, Ann, Douglas M. Sloane, and Linda H. Aiken. 2013. 'An Increase In The Number Of Nurses With Baccalaureate Degrees Is Linked To Lower Rates Of Postsurgery Mortality', *Health Affairs*, 32: 579-86.

Lake, Eileen T. 2002. 'Development of the practice environment scale of the nursing work index†‡', *Research in nursing & health*, 25: 176-88.

Lancet-Commission. 1932. 'The Lancet Commission on Nursing (Appointed 1930)'. http://jamanetwork.com/journals/jama/article-abstract/280126.

Lankshear, Annette J, Trevor A Sheldon, and Alan Maynard. 2005. 'Nurse staffing and healthcare outcomes: a systematic review of the international research evidence', *Advances in Nursing Science*, 28: 163-74.

Laschinger, Heather K Spence, and Michael P Leiter. 2006. 'The impact of nursing work environments on patient safety outcomes: The mediating role of burnout engagement', *Journal of Nursing Administration*, 36: 259-67.

Leininger, Madeline M, and Maggie Phoenix Gilead. 1984. 'Care: The Essence of Nursing and Health', *Home Healthcare Now*, 2: 10.

Levine, Eugene. 1961. 'Nurse staffing in hospitals', *AJN The American Journal of Nursing*, 61: 65&hyhen.

Li, Yu-Fang, Eileen T Lake, Anne E Sales, Nancy D Sharp, Gwendolyn T Greiner, Elliott Lowy, Chuan-Fen Liu, Pamela H Mitchell, and Julie A Sochalski. 2007. 'Measuring nurses' practice environments with the revised nursing work index: evidence from registered nurses in the Veterans Health Administration', *Research in Nursing & Health*, 30: 31-44.

Lindqvist, Rikard, Lisa Smeds Alenius, Sara Runesdotter, Anneli Ensio, Virpi Jylhä, Juha Kinnunen, Ingeborg Strømseng Sjetne, Christine Tvedt, Maria Wiberg Tjønnfjord, and Carol Tishelman. 2014. 'Organization of nursing care in three Nordic countries: relationships between nurses' workload, level of involvement in direct patient care, job satisfaction, and intention to leave', *BMC Nursing*, 13: 27.

Lindqvist, Rikard, Lisa Smeds Alenius, Peter Griffiths, Sara Runesdotter, and Carol Tishelman. 2015. 'Structural characteristics of hospitals and nurse-reported care quality, work environment, burnout and leaving intentions', *Journal of Nursing Management*, 23: 263-74.

Macchiarini-Case-Investigation. 2016. "The Macchiarini Case: Investigation of the synthetic trachea transplantations at Karolinska University Hospital."

Mackenbach, JP, MH Bouvier-Colle, and E Jougla. 1990. '" Avoidable" mortality and health services: a review of aggregate data studies', *Journal of Epidemiology and Community Health*, 44: 106-11.

Makary, Martin A, and Michael Daniel. 2016. 'Medical error—the third leading cause of death in the US', *BMJ*, 353: i2139.

Mark, Barbara A. 1995. 'The black box of patient outcomes research', *Image: the Journal of Nursing Scholarship*, 27: 42-42.

Mark, Barbara A, Linda C Hughes, and Cheryl Bland Jones. 2004. 'The role of theory in improving patient safety and quality health care', *Nursing Outlook*, 52: 11-16.

McGillis-Hall, L, Diane Doran, and George H Pink. 2004. 'Nurse staffing models, nursing hours, and patient safety outcomes', *Journal of Nursing Administration*, 34: 41-45.

McHugh, Matthew D, Lesly A Kelly, Herbert L Smith, Evan S Wu, Jill M Vanak, and Linda H Aiken. 2013. 'Lower mortality in magnet hospitals', *The Journal of Nursing Administration*, 43: S4.

McHugh, Matthew D, and Amy Witkoski Stimpfel. 2012. 'Nurse reported quality of care: a measure of hospital quality', *Research in Nursing & Health*, 35: 566-75.

McKee, Martin, Jennifer Dixon, and Laurent Chenet. 1994. 'Making routine data adequate to support clinical audit', *BMJ: British Medical Journal*, 309: 1246.

Merrifield, N. 2016. "Nursing associate role will 'supplement, not replace' RNs." In *Nursing Times*. Harrow.

Michie, S, and S Williams. 2003. 'Reducing work related psychological ill health and sickness absence: a systematic literature review', *Occupational & Environmental Medicine*, 60: 3-9.

Needleman, J. 2016. 'The Economic Case for Fundamental Nursing Care', *Nursing Leadership (Toronto, Ont.)*, 29: 26.

Needleman, Jack, Peter I Buerhaus, Maureen Stewart, Katya Zelevinsky, and Soeren Mattke. 2006. 'Nurse staffing in hospitals: is there a business case for quality?', *Health Affairs*, 25: 204-11.

Needleman, Jack, Peter Buerhaus, Soeren Mattke, Maureen Stewart, and Katya Zelevinsky. 2002. 'Nurse-staffing levels and the quality of care in hospitals', *New England Journal of Medicine*, 346: 1715-22.

Needleman, Jack, Peter Buerhaus, V Shane Pankratz, Cynthia L Leibson, Susanna R Stevens, and Marcelline Harris. 2011. 'Nurse staffing and inpatient hospital mortality', *New England Journal of Medicine*, 364: 1037-45.

NHS-England. 2013. 'Our Ambition to Reduce Premature Mortality: A resource to support commissioners in setting a level of ambition'. London.

NICE. 2014. "Safe staffing for nursing in adult inpatient wards in acute hospitals (SG1)." In. London: National Institute for Health and Care Excellence

Nightingale, Florence. 1858. *Notes on Matters Affecting the Health, Efficiency, and Hospital Administration of the British Army: Founded Chiefly on the Experience of the Late War. By Florence Nightingale. Presented by Request to the Secretary of State for War* (Harrison and Sons, St. Martin's Lane, WC).

Nolte, Ellen, and Martin McKee. 2004. *Does health care save lives? Avoidable mortality revisited* . The Nuffield Trust. London.

Nursing-Times. 2013. "NICE to start work on safe staffing guideline." 13 Nov 2013. *Nursing Times*.

O'Brien–Pallas, Linda, X Mingyang Li, Sping Wang, Raquel M Meyer, and Donna Thomson. 2010. 'Evaluation of a patient care delivery model: System outcomes in acute cardiac care', *Canadian Journal of Nursing Research*, 42: 98-120.

Pearl, Judea, and Thomas S Verma. 1995. 'A theory of inferred causation', *Studies in Logic and the Foundations of Mathematics*, 134: 789-811.

Pearson, Alan. 1988. *Primary nursing: nursing in the Burford and Oxford nursing development units* . Croom Helm.

Prytherch, David R, Gary B Smith, Paul E Schmidt, and Peter I Featherstone. 2010. 'ViEWS—towards a national early warning score for detecting adult inpatient deterioration', *Resuscitation*, 81: 932-37.

Rafferty, A. M., S. P. Clarke, J. Coles, J. Ball, P. James, M. McKee, and L. H. Aiken. 2007. 'Outcomes of variation in hospital nurse staffing in English hospitals: cross-sectional analysis of survey data and discharge records', *International Journal of Nursing Studies*, 44: 175-82.

RCSE. 2011. "The Higher Risk General Surgical Patient: towards improved care for a forgotten group." In. London: Royal College of Surgeons England.

Reason, J. 1998. 'How necessary steps in a task get omitted: revising old ideas to combat a persistent problem', *Cognitive Technol*, 3: 24-32.

Reason, James. 1990. *Human error* .Cambridge University Press.

Richardson, Gerald, Alan Maynard, Nicky Cullum, and David Kindig. 1998. 'Skill mix changes: substitution or service development?', *Health Policy*, 45: 119-32.

Ridelberg, Mikaela, Kerstin Roback, and Per Nilsen. 2014. 'Facilitators and barriers influencing patient safety in Swedish hospitals: a qualitative study of nurses' perceptions', *BMC Nursing*, 13: 1.

Rutstein , David D., William Berenberg , Thomas C. Chalmers , Charles G. 3rd Child , Alfred P. Fishman , Edward B. Perrin , Jacob J. Feldman , Paul E. Leaverton , J. Michael Lane , David J. Sencer , and Carleton C. Evans 1976. 'Measuring the Quality of Medical Care', *New England Journal of Medicine*, 294: 582-88.

Sales, A., N. Sharp, Y. F. Li, E. Lowy, G. Greiner, C. F. Liu, A. Alt-White, C. Rick, J. Sochalski, P. H. Mitchell, G. Rosenthal, C. Stetler, P. Cournoyer, and J. Needleman. 2008. 'The association between nursing factors and patient mortality in the Veterans Health Administration: the view from the nursing unit level', *Medical Care*, 46: 938-45.

Sales, Anne. 2015. 'Nurse staffing matters: now what?', *BMJ Quality & Safety*, 24: 241-43.

Sasichay-Akkadechanunt, T., C.C. Scalzi, and A.F. Jawad. 2003. 'The relationship between nurse staffing and patient outcomes', *Journal of Nursing Administration*, 33: 478.

Schubert, M., S. P. Clarke, T. R. Glass, B. Schaffert-Witvliet, and S. De Geest. 2009. 'Identifying thresholds for relationships between impacts of rationing of nursing

care and nurse- and patient-reported outcomes in Swiss hospitals: a correlational study', *International Journal of Nursing Studies*, 46: 884-93.

Schubert, Maria, Sean P Clarke, Linda H Aiken, and Sabina De Geest. 2012. 'Associations between rationing of nursing care and inpatient mortality in Swiss hospitals', *International Journal for Quality in Health Care*: mzs009.

Schubert, Maria, Tracy R Glass, Sean P Clarke, Bianca Schaffert-Witvliet, and Sabina De Geest. 2007. 'Validation of the Basel extent of rationing of nursing care instrument', *Nursing Research*, 56: 416-24.

Scott, Joan Gleason, Julie Sochalski, and Linda Aiken. 1999. 'Review of magnet hospital research: findings and implications for professional nursing practice', *Journal of Nursing Administration*, 29: 9-19.

Sellick, Kenneth J, Shirley Russell, and Jacqueline L Beckmann. 1983. 'Primary nursing: an evaluation of its effects on patient perception of care and staff satisfaction', *International Journal of Nursing Studies*, 20: 265-73.

Sermeus, Walter, Linda H Aiken, Koen Van den Heede, Anne M Rafferty, Peter Griffiths, Maria T Moreno-Casbas, Reinhard Busse, Rikard Lindqvist, Anne P Scott, and Luk Bruyneel. 2011. 'Nurse forecasting in Europe (RN4CAST): rationale, design and methodology', *BMC Nursing*, 10: 6.

Shekelle, Paul G. 2013a. 'Effect of Nurse-to-Patient Staffing Ratios on Patient Morbidity and Mortality.' in, *Making Health Care Safer II: An Updated Critical Analysis of the Evidence for Patient Safety Practices.* (Agency for Healthcare Research and Quality US).

Shekelle. P 2013b. 'Nurse–patient ratios as a patient safety strategy: a systematic review', *Annals of Internal Medicine*, 158: 404-09.

Sheward, L., J. Hunt, S. Hagen, M. Macleod, and J. Ball. 2005. 'The relationship between UK hospital nurse staffing and emotional exhaustion and job dissatisfaction', *Journal of Nursing Management*, 13: 51-60.

Shojania, Kaveh G, and Mary Dixon-Woods. 2016. 'Estimating deaths due to medical error: the ongoing controversy and why it matters', *BMJ Quality & Safety*: bmjqs-2016-006144.

Shojania, Kaveh G, and Alan J Forster. 2008. 'Hospital mortality: when failure is not a good measure of success', *Canadian Medical Association Journal*, 179: 153-57.

Silber, Jeffrey H, Sean K Kennedy, Orit Even-Shoshan, Wei Chen, Laurie F Koziol, Ann M Showan, and David E Longnecker. 2000. 'Anesthesiologist direction and patient outcomes', *The Journal of the American Society of Anesthesiologists*, 93: 152-63.

Silber, Jeffrey H, and Paul R Rosenbaum. 1997. 'A spurious correlation between hospital mortality and complication rates: the importance of severity adjustment', *Medical Care*: OS77-OS92.

Silber, Jeffrey H, Sankey V Williams, Henry Krakauer, and J Sanford Schwartz. 1992. 'Hospital and patient characteristics associated with death after surgery: a study of adverse occurrence and failure to rescue', *Medical Care*: 615-29.

Smeds-Alenius, Lisa, Carol Tishelman, Rikard Lindqvist, Sara Runesdotter, and Matthew D McHugh. 2016. 'RN assessments of excellent quality of care and

patient safety are associated with significantly lower odds of 30-day inpatient mortality: A national cross-sectional study of acute-care hospitals', *International Journal of Nursing Studies*, 61: 117-24.

Sochalski, J. 2004. 'Is more better?: the relationship between nurse staffing and the quality of nursing care in hospitals', *Medical Care*, 42: II-67-73.

Sochalski, J., R. T. Konetzka, J. Zhu, and K. Volpp. 2008. 'Will mandated minimum nurse staffing ratios lead to better patient outcomes?', *Medical Care*, 46: 606-13.

Soop, M., U. Fryksmark, M. Köster, and B. Haglund. 2009. 'The incidence of adverse events in Swedish hospitals: a retrospective medical record review study', *International Journal of Quality Health Care*, 21.

Spilsbury, Karen, Joy Adamson, Karl Atkin, Karen Bloor, Roy Carr-Hill1, Dorothy McCaughan, Hugh McKenna, and Ann Wakefield. 2011. 'Challenges and opportunities associated with the introduction of assistant practitioners supporting the work of registered nurses in NHS acute hospital trusts in England', *Journal of Health Services Research & Policy*, 16: 50-56.

Squires, Allison, Linda H Aiken, Koen van den Heede, Walter Sermeus, Luk Bruyneel, Rikard Lindqvist, Lisette Schoonhoven, Ingeborg Stromseng, Reinhard Busse, and Tomasz Brzostek. 2013. 'A systematic survey instrument translation process for multi-country, comparative health workforce studies', *International Journal of Nursing Studies*, 50: 264-73.

Swedish-Accident-Investigation-Authority. 2013. "Statens haverikommission: Dödsfall på hjärtklinikens hjärtintensivavdelning N25 vid Karolinska Universitetssjukhuset i Solna, Stockholms län, den 17 oktober 2010. (Slutrapport RO 2013:02)

Swedish Association of Local Authorities and Regions. 2011. "Sjukvårdsdata i Fokus." Sveriges Kommuner och Landsting.

Thornlow, Deirdre K, Ruth Anderson, and Eugene Oddone. 2009. 'Cascade iatrogenesis: factors leading to the development of adverse events in hospitalized older adults', *International Journal of Nursing Studies*, 46: 1528-35.

Tourangeau, Ann E. 2005. 'A theoretical model of the determinants of mortality', *Advances in Nursing Science*, 28: 58-69.

Tourangeau, Ann E. 2010. 'Mortality rate as a nurse-sensitive outcome.' in, *Nursing Outcomes: The state of science* (Jones & Bartlett: Sudbury).

Van den Heede, Koen, Luwis Diya, Emmanuel Lesaffre, Arthur Vleugels, and Walter Sermeus. 2008. 'Benchmarking nurse staffing levels: the development of a nationwide feedback tool', *Journal of Advanced Nursing*, 63: 607-18.

Vincent, Charles, Krishna Moorthy, Sudip K Sarker, Avril Chang, and Ara W Darzi. 2004. 'Systems approaches to surgical quality and safety: from concept to measurement', *Annals of Surgery*, 239: 475-82.

Waterson, Patrick, P Griffiths, C Stride, J Murphy, and Sue Hignett. 2010. 'Psychometric properties of the Hospital Survey on Patient Safety Culture: findings from the UK', *Quality and Safety in Health Care*, 19: e2-e2.

West, Elizabeth, David N Barron, David Harrison, Anne Marie Rafferty, Kathy Rowan, and Colin Sanderson. 2014. 'Nurse staffing, medical staffing and mortality in

intensive care: an observational study', *International Journal of Nursing Studies*, 51: 781-94.

Whelan, J. 2000. 'Where Did All the Nurses Go? (Based on PhD. diss, University of Pennsylvania, 2000) ', Accessed 17th Feb 2017. http://www.nursing.upenn.edu/nhhc/workforce-issues/where-did-all-the-nurses-go/.

WHO. 2016. " Global Strategic Directions for Strengthening Nursing and Midwifery 2016-2020." Geneva.

You, Li-ming, Linda H. Aiken, Douglas M. Sloane, Ke Liu, Guo-ping He, Yan Hu, Xiao-lian Jiang, Xiao-han Li, Xiao-mei Li, Hua-ping Liu, Shao-mei Shang, Ann Kutney-Lee, and Walter Sermeus. 2013. 'Hospital nursing, care quality, and patient satisfaction: Cross-sectional surveys of nurses and patients in hospitals in China and Europe', *International Journal of Nursing Studies*, 50: 154-61.

Zúñiga, F., D. Ausserhofer, J. P. H. Hamers, S. Engberg, M. Simon, and R. Schwendimann. 2015. 'Are Staffing, Work Environment, Work Stressors, and Rationing of Care Related to Care Workers' Perception of Quality of Care? A Cross-Sectional Study', *Journal of the American Medical Directors Association*, 16: 860-66.

APPENDIX

A1. Reviewed studies: design and validity

A2. Study I additional data: Model results for RN staffing and individual

A1. Nurse staffing & patient outcome studies: design and validity

	Design type	Internal validity	External validity	Strong/ medium on both
1. Ausserhofer 2013	CS	-	++	
2. Ball et al 2013	CS	+	++	✓
3. Blegen 2011	CS	++	+	✓
4. Blegen & Goode 1998	CS/RO	-	-	
5. Blegen & Vaughn 1998	RO	+	+	
6. Chang 2011	CS	-	++	
7. Cho 2003	RO	++	+	✓
8. Donaldson 2005	RO	+	++	✓
9. Duffield 2011	CS/RO	-	+	
10. Estabrooks 2005	CS	++	+	✓
11. Frith 2010	CS	-	-	
12. Frith 2012	RO	+	-	
13. Hart 2011	CS/RO	-	-	
14. He 2013	CS	++	++	✓✓
15. Ibe 2008	CS/RO	-	+	
16. Kutney Lee 2013	RO	++	+	✓
17. Lake 2010	CS	+	+	
18. Manojlvich 2011	RO	-	-	
19. McGillis Hall 2004	CS	-	+	
20. Needleman 2011	RO	++	+	✓
21. O'Brien Pallas 2010	PO	+	-	
22. O'Brien Pallas 2010 b	CS	-	-	
23. Park 2012	CS/RO	++	+	✓
24. Patrician 2011	RO	++	++	✓✓
25. Potter 2003	PO	+	-	
26. Sales 2008	CS/RO	++	++	✓✓
27. Seago 2006	RO	-	-	
28. Shever 2005	RO	-	-	
29. Sochalski 2008	CS/RO	++	+	✓
30. Spetz 2013	CS/RO	++	++	✓✓
31. Staggs 2012	RO	+	+	
32. Tschannen 2010	CS	+	++	✓
33. Twigg 2013	RO	++	+	✓
34. Unruh 2007	RO	-	-	
35. Weiss2011	CS	+	-	

Source: GRIFFITHS, P., BALL, J., DRENNAN, J., JAMES, L., JONES, J., RECIO, A. & SIMON, M. 2014.

'The association between patient safety outcomes and nurse/health care assistant skill mix and staffing levels and factors that may influence staffing requirements.' University of Southampton.

A2: Model results for RN staffing and individual aspects of missed care (Study I additional data)

Patients per Registered Nurse																	
	up to 6.13				6.14 - 7.33				7.40 - 9.25				9.33 - 11.50				
	OR	L95	U95	p	OR	L95	U95	p	OR	L95	U95	p	OR	L95	U95	Sdf^2	p
Adequate patient surveillance	0.39	0.29	0.54	§	0.58	0.43	0.79	§	0.72	0.54	0.97	†	0.80	0.61	1.06	38.15	<.001
Adequately document nursing care	0.58	0.41	0.81	‡	1.01	0.74	1.40		1.23	0.90	1.67		0.98	0.74	1.31	26.34	<.001
Comfort/talk with patients	0.49	0.36	0.67	§	0.66	0.48	0.91	†	0.82	0.60	1.11		0.86	0.65	1.15	25.71	<.001
Skin care	0.47	0.32	0.68	§	0.77	0.54	1.09		0.80	0.57	1.12		0.80	0.59	1.10	17.04	0.002
Educating patients & family	0.70	0.52	0.94	†	0.84	0.62	1.13		0.96	0.72	1.28		1.18	0.91	1.55	15.72	0.003
Develop/update nursing care plans/pathways	0.63	0.46	0.85	‡	0.95	0.70	1.29		0.91	0.68	1.22		1.00	0.76	1.31	14.74	0.005
Oral hygiene [1]	0.62	0.45	0.87	‡	0.86	0.62	1.18		0.96	0.70	1.29		0.90	0.67	1.19	10.49	0.033
Planning care	0.63	0.45	0.88	‡	0.85	0.61	1.18		0.87	0.63	1.19		0.99	0.74	1.32	9.82	0.044
Treatments & procedures	0.54	0.34	0.87	†	0.75	0.48	1.17		0.97	0.63	1.47		0.91	0.61	1.35	9.24	0.055
Frequent changing of patient position	0.71	0.51	1.00		1.01	0.73	1.39		0.96	0.70	1.31		0.97	0.72	1.29	6.68	0.154
Administer medications on time	0.66	0.46	0.94	†	0.89	0.63	1.26		0.90	0.64	1.25		0.84	0.62	1.15	6.14	0.189
Pain management	1.49	0.81	2.73		1.59	0.88	2.88		1.60	0.90	2.86		1.65	0.97	2.81	3.85	0.426
Prepare patients & families for discharge	0.87	0.61	1.24		0.92	0.65	1.31		0.96	0.68	1.35		0.99	0.72	1.35	0.77	0.943

[1]
Odds ratios (OR) from main effects model (without the patients per RN x patients per HCA interaction).
[2] 11.67 and over was the reference category for the calculation of ORs; OR for this category equals 1.

† < .05; ‡ < .01; § < .001